

178655

STIC-Biotech/ChemLib

From: Chan, Christina
Sent: Monday, February 06, 2006 10:21 AM
To: Yu, Misook; STIC-Biotech/ChemLib
Subject: RE: rsuh search request 10/111,257

~~Please rush. Thanks Chris~~

CRIF

Chris Chan
TC 1600 New Hire Training Coordinator and SPE 1644
(571)-272-0841
Remsen, 3E89

-----Original Message-----

From: Yu, Misook
Sent: Monday, February 06, 2006 9:09 AM
To: Chan, Christina
Subject: rsuh search request 10/111,257

Pls approve rush search for the case due this biweek.

Stic, pls search SEQ ID NO: 3 (nucleic acid).

Examiner Misook Yu, Ph.D.
571-272-0839 (phone)
571-273-0839 (fax)
Art Unit 1642
REM-3D29 (Office)
REM-3C18 (Mail Box)
400 Dulany Street
Alexandria, VA 22314

Searcher: _____
Searcher Phone: _____
Date Searcher Picked up: _____
Date completed: _____
Searcher Prep Time: _____
Online Time: _____

Type of Search
NA# _____ AA# _____
S/L: _____ Oligomer: _____
Encode/Transl: _____
Structure #: _____ Text: _____
Inventor: _____ Litigation: _____

Vendors and cost where applicable
STN: _____
DIALOG: _____
QUESTEL/ORBIT: _____
LEXIS/NEXIS: _____
SEQUENCE SYSTEM: _____
WWW/Internet: _____
Other (Specify): _____

THIS PAGE BLANK (USPTO)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 6, 2006, 16:26:26 ; Search time 8168 Seconds
(without alignments) 11542.112 Million cell updates/sec

Title: US-10-111-257-3

Perfect score: 2015

Sequence: 1 cgcctcgcacgcggtgacgc.....gtgaagttctcaactcaa 2015

Scoring table: IDENTITY_NUC

Gapop 10.0, Gapext 1.0

Searched: 41078325 seqs, 23393541228 residues

82156650

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

Listing first 45 summaries

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1964.2	97.5	5165	4	CR749560 Homo sapi
2	1845.6	91.6	3850	4	CR857633 Pongo pyg
3	1383	68.6	1383	11	DO045794 Homo sapi
4	867	43.0	935	5	BU944114 AGENCOURT
5	802.2	39.8	839	3	BM014389 603640139
6	748	37.1	909	7	CT001323 CT001323
7	738.4	36.6	915	5	BU176917 AGENCOURT
8	730.6	36.3	940	5	BU539183 AGENCOURT
9	718.2	35.6	756	1	AM736896 IL3-CT021
10	711.6	35.3	890	5	BO719354 AGENCOURT
11	707	35.1	731	8	CR869347 HESG4_34
12	666	33.1	692	7	CN389062 170004240
13	664	33.0	696	7	CN389072 170004240
14	663.2	32.9	709	11	DO045795 Pan trogl
15	648.2	32.2	656	7	CK003957 AGENCOURT
16	647.4	32.1	684	7	CN389071 170006000
17	647.2	32.1	684	1	AM959984 EST372055
18	618.6	30.7	1755	11	DO050311 Homo sapi
19	611	30.3	680	2	BE742835 601574472
20	604.8	30.0	873	5	BF685522 602142247
21	592.4	29.4	715	1	AM601306 PMO-BT034

23	587.8	29.2	658	2	BF001707 7929a12.x
24	582.2	28.9	601	1	AM367440 MRO-HT016
25	581.8	28.9	840	5	BU176890 AGENCOURT
26	570	28.3	3261	4	CR860800 Pongo pyg
27	567	28.1	582	3	BP293039 BP293039
28	566.6	28.1	697	6	CD641275 AGENCOURT
29	564.4	28.0	2919	4	BC040591 Homo sapi
30	562.8	27.9	2299	4	AF161544 Homo sapi
31	562.8	27.9	4132	4	HSM805700 Homo sapi
32	558.2	27.7	617	8	W72665
33	553	27.4	702	10	AG142996 Pan trogl
34	546.6	27.1	594	2	BE178353 RC3-HT060
35	546	27.1	1467	11	DO032599 Homo sapi
36	543.8	27.0	616	2	BE178002 RC3-HT060
37	541.6	26.9	744	10	AG143038 Pan trogl
38	538.4	26.7	744	6	CD521516 AGENCOURT
39	538	26.7	1929	11	DO045890 Homo sapi
40	534.2	26.5	548	9	AO580696 RPI-11-4
41	534.2	26.5	679	2	BG718400 602696429
42	531.4	26.4	755	6	CD520335 AGENCOURT
43	529.8	26.3	1694	11	DO048711 Homo sapi
44	529	26.3	582	3	BP361237 BP361237
45	528.6	26.2	612	1	AM749131 PMO-BT034

ALIGNMENTS

RESULT 1
LOCUS CR749560 5165 bp mRNA linear HTC CR749560
DEFINITION Homo sapiens mRNA; cDNA DKFP686L0787 (from clone DKFP686L0787).
ACCESSION CR749560.1 GI:51476788
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.
1 (bases 1 to 5165)
Poustka, A., Albert, R., Moosmayer, P., Schnupp, I., Wellenreuther, R.,
Mewes, H.W., Well, B., Amid, C., Osanger, A., Fobo, G., Han, M. and
Wiemann, S.
The German cDNA Consortium
Direct Submission
Submitted (17-AUG-2004) MIPS, Ingolstaedter Landstr.1, D-85764
Neuherberg, GERMANY
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ), Email s.wiemann@dkfz-heidelberg.de;
Sequenced by DKFZ (German Cancer Research Center,
Heidelberg/Germany) within the cDNA sequencing consortium of the
German Genome Project. This clone (DKFP686L0787) is available at
the RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH in
Berlin, Germany. Please contact RZPD for ordering:
http://www.rzpd.de/cgi-bin/products/ci.cgi?CloneID=DKFP686L0787
Further information about the clone and the sequencing project is
available at http://mips.gsf.de/projects/cdna/.

COMMENT
JOURNAL
CONSRTM
TITLE
JOURNAL

FEATURES
source
1..5165
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="DKFP686L0787"
/tissue_type="retina"
/clone_lib="686 (synonym: hicc3). Vector pSport1_Sfi; host
DH10B; sites SfiIa + SfiIIB"
/dev_stage="adult"
/note="zinc finger protein 331"
1..5165
/gene="DKFP686L0787"
1345..2736
/gene="DKFP686L0787"

gene
CDs
AM601306 PMO-BT034

```
/codon_start=1  
/product="hypothetical protein"  
/protein_id="CAH18357.1"  
/db_xref="GI:51476789"  
/translation="MAQGLVTPADVAIDPDSQSEWACINSARDLYMDVLEENSLVLS  
LDLESAYENKSLPTEKNLHEIRASRNSRKSJGRNMTCEGTLERPQRSGRYVNO  
MILNVKQPRATREQTPRTHQPHHKENSEPCDQKAFRQGLQLOHOKIHGKEPYE  
CECKKAFRMQNLTOHOKIHGKEPYECKDQKAFRMSSLVIKRITHGKEPYE  
DCGKAFRDELTOHORPHITGKEPYECKDQKAFRMSSLVYLIOHKITHGKEPYE  
GKAFI GSSLIQHKRITHGKEPYEEOCKDQKAFRMSSLVYLIOHKITHGKEPYE  
AFRMSSSLVHERITHGKEPYECKDQKAFRMSSLVYLIOHKITHGKEPYE  
IYSSSLVHERITHGKEPYECKDQKAFRMSSLVYLIOHKITHGKEPYE  
LNHLREHORIHNS"
```

ORIGIN

```
Query Match 97.5%; Score 1964.2; DB 4; Length 5165;  
Best Local Similarity 99.8%; Pred. No. 0;  
Matches 1966; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
  
OY 47 CAGGCCAGCATCCTTCAGAAAAAGCATCCCGAGAGAGAGAGAACTGTTAAACATCTT 106  
DB 1205 CAGGCCAGCATCCTTCAGAAAAAGCATCCCGAGAGAGAGAGAACTGTTAAACATCTT 1264  
OY 107 AGGTGAGCTTACCTCTCGGAAATTTGTCTTTCAGTGGAAACCCCGAGAGAGAGAGAG 166  
DB 1265 AGGTGAGCTTACCTCTCGGAAATTTGTCTTTCAGTGGAAACCCCGAGAGAGAGAGAG 1324  
OY 167 AGTTCCTGAGTCTTAAACAAATGGCCAGAGGTTGGTGAAGTTGGCCGAGAGTACCATAG 226  
DB 1325 AGTTCCTGAGTCTTAAACAAATGGCCAGAGGTTGGTGAAGTTGGCCGAGAGTACCATAG 1384  
OY 227 ACTTTTCTCAGAGAGAGAGGAGGCTGTCTGAACCTCTGCTCAGAGAGAGAGAGAGAGAG 286  
DB 1385 ACTTTTCTCAGAGAGAGAGGAGGCTGTCTGAACCTCTGCTCAGAGAGAGAGAGAGAGAG 1444  
OY 287 TGATGCTGAGAACTTACAGTAACTTGGTCTCACTGAGTTTGGAGTACATATGAAATAA 346  
DB 1445 TGATGCTGAGAACTTACAGTAACTTGGTCTCACTGAGTTTGGAGTACATATGAAATAA 1504  
OY 347 AGAGTTTACTCTACAGAAAAAAACATTCAGGAATTAAGAGGCTTCCAAAGGAAATTCAGAT 406  
DB 1505 AGAGTTTACTCTACAGAAAAAAACATTCAGGAATTAAGAGGCTTCCAAAGGAAATTCAGAT 1564  
OY 407 GAAGAAGTAATCCCTTGGCCGTAATCTGATATGTGAAGTACGCTTGAAGAAGAGAGAG 466  
DB 1565 GAAGAAGTAATCCCTTGGCCGTAATCTGATATGTGAAGTACGCTTGAAGAAGAGAGAG 1624  
OY 467 GCTCCAGAGGAGAGTATGTCAATCAATGATCATCAATTAATGTCAAAAGGCTGTACTA 526  
DB 1625 GCTCCAGAGGAGAGTATGTCAATCAATGATCATCAATTAATGTCAAAAGGCTGTACTA 1684  
OY 527 GGAAGAAGAGAGGCTTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 586  
DB 1685 GGAAGAAGAGAGGCTTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1744  
OY 587 AGGAGCTGGGAGAGAGGCTTGTAGTCTGGCTATCAACTTACTGTAACATCAGAAAAATCAT 646  
DB 1745 AGGAGCTGGGAGAGAGGCTTGTAGTCTGGCTATCAACTTACTGTAACATCAGAAAAATCAT 1804  
OY 647 CTGCTGAGAAACCTTATGAAATGTAAAGATGTAAAGAGGCTTCCGTTGGGGGCAATCAGC 706  
DB 1805 CTGCTGAGAAACCTTATGAAATGTAAAGATGTAAAGAGGCTTCCGTTGGGGGCAATCAGC 1864  
OY 707 TTAATCAACATCAAAATAATCATACGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 766  
DB 1865 TTAATCAACATCAAAATAATCATACGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1924  
OY 767 AGGCTTTGATGAGGAGGCTCAAGGCTGTTATTCATTAAGAGATTCATCTGTGAAAGAAC 826  
DB 1925 AGGCTTTGATGAGGAGGCTCAAGGCTGTTATTCATTAAGAGATTCATCTGTGAAAGAAC 1984  
OY 827 CCTATGAATGTAAAGACTGTGAAAGAGGCTTTGGCGGTGTGATGAGCTCACTCAGACACC 886  
DB 827 CCTATGAATGTAAAGACTGTGAAAGAGGCTTTGGCGGTGTGATGAGCTCACTCAGACACC 886
```

```
DB 1985 CCTATGAATGTAAAGACTGTGAAAGAGGCTTTGGCGGTGTGATGAGCTCACTCAGACACC 2044  
OY 887 AGAGATTCACACACTGGGAGAGAAAGACTACGAATGCAAAAGACTGTGGAGAACCTTTAGCC 946  
DB 2045 AGAGATTCACACACTGGGAGAGAAAGACTACGAATGCAAAAGACTGTGGAGAACCTTTAGCC 2104  
OY 947 GTGTGTATTAACCTTATTCAGCAACAAGAAATTCATATGTGGGAGAGAGGCTTTACGAGTGT 1006  
DB 2105 GTGTGTATTAACCTTATTCAGCAACAAGAAATTCATATGTGGGAGAGAGGCTTTACGAGTGT 2164  
OY 1007 AAGACTGTGGGAGAGGCTTTATTTGTGTGCTCAAGGCTCAATTCAGATTAAGAAATTCACA 1066  
DB 2165 AAGACTGTGGGAGAGGCTTTATTTGTGTGCTCAAGGCTCAATTCAGATTAAGAAATTCACA 2224  
OY 1067 CAGTGAAGAAACCTTATGAAATGTCAAGATGTGGAGAGGCTTTTACTCGAGTCAATTAAC 1126  
DB 2225 CAGTGAAGAAACCTTATGAAATGTCAAGATGTGGAGAGGCTTTTACTCGAGTCAATTAAC 2284  
OY 1127 TTAATCAAGATCAAGAAAGATCCACACCGGTGAGAGGCTTCAAGATGTAAAGAGTGTGG 1186  
DB 2285 TTAATCAAGATCAAGAAAGATCCACACCGGTGAGAGGCTTCAAGATGTAAAGAGTGTGG 2344  
OY 1187 AGGCTTTGCTGGGGGTTGAGGCTCGTTAAGCAGAGAGATACATACGAGGAGAGAGC 1246  
DB 2345 AGGCTTTGCTGGGGGTTGAGGCTCGTTAAGCAGAGAGATACATACGAGGAGAGAGC 2404  
OY 1247 CGTACAGTGCACAGAAATGTGGAAAGGCTTCAATTTGTGCTATCACTCAGACAG 1306  
DB 2405 CGTACAGTGCACAGAAATGTGGAAAGGCTTCAATTTGTGCTATCACTCAGACAG 2464  
OY 1307 AAGAAATCCACACAGGCGAAACCCCGTATTAATGTAAAGAGTGTGGAAAGGCTTTCATT 1366  
DB 2465 AAGAAATCCACACAGGCGAAACCCCGTATTAATGTAAAGAGTGTGGAAAGGCTTTCATT 2524  
OY 1367 ATGGAATCGAGGCTCGTGAAGAGAGAGAAATTCATACCGGGGTGAACCCCTAATGGGTGA 1426  
DB 2525 ATGGAATCGAGGCTCGTGAAGAGAGAGAAATTCATACCGGGGTGAACCCCTAATGGGTGA 2584  
OY 1427 CAGAAATGTGGAGAGAGCTTTTACTACGAGCATCAGCTTACACAAATCAGAAAAAGCACA 1486  
DB 2585 CAGAAATGTGGAGAGAGCTTTTACTACGAGCATCAGCTTACACAAATCAGAAAAAGCACA 2644  
OY 1487 GTGGGGCGAAATCTTACGAATGTAAAGAGTGTGGAGAGAGAGAGAGAGAGAGAGAG 1546  
DB 2645 GTGGGGCGAAATCTTACGAATGTAAAGAGTGTGGAGAGAGAGAGAGAGAGAGAGAG 2704  
OY 1547 TCCGAGAAATCAAGAGAGATCCACAAAGTTGAAGAGGCTTTGAAGAGAGAGAGAGAG 1606  
DB 2705 TCCGAGAAATCAAGAGAGATCCACAAAGTTGAAGAGGCTTTGAAGAGAGAGAGAGAG 2764  
OY 1607 GTATCTATGTTTGGCTTCCACAGTTTGTACTGCACTGCACTGCACTGCACTGCACTGCA 1666  
DB 2765 GTATCTATGTTTGGCTTCCACAGTTTGTACTGCACTGCACTGCACTGCACTGCACTGCA 2824  
OY 1667 TAAATGAAATAATTCAGAAATAAAGATTTTAAGTCTCAATGTGTGTGCTTCTGAGAT 1726  
DB 2825 TAAATGAAATAATTCAGAAATAAAGATTTTAAGTCTCAATGTGTGTGCTTCTGAGAT 2884  
OY 1727 GCGTGAATGAATCTCGCGGTGCTCCGCTCCAGCGGCGGAGAGAGAGAGAGAGAGAGAG 1786  
DB 2885 GCGTGAATGAATCTCGCGGTGCTCCGCTCCAGCGGCGGAGAGAGAGAGAGAGAGAGAG 2944  
OY 1787 TCCAGACATCCACGCTGTATACGCCACCACTGCTGTAAGTACTTAAAGAGCGGTCTTGG 1846  
DB 2945 TCCAGACATCCACGCTGTATACGCCACCACTGCTGTAAGTACTTAAAGAGCGGTCTTGG 3004  
OY 1847 TGATCAGATCAACTATCCAGATCAACAGTCTGTGCTGCCAAGTACTCTCACTTTGCTT 1906  
DB 3005 TGATCAGATCAACTATCCAGATCAACAGTCTGTGCTGCCAAGTACTCTCACTTTGCTT 3064  
OY 1907 AACAGTGGCCCGCAGAGAGAGAGAGAGATGATGCTGTGATTTGCGAATATGCGAAAGAGAG 1966  
DB 3065 AACAGTGGCCCGCAGAGAGAGAGAGATGATGCTGTGATTTGCGAATATGCGAAAGAGAG 3124
```


QY 1967 CCACGAAGTCTCTCTTTAAATGAAGTGAAGTCTCAACTTAA 2015
DB 3125 CCACGAAGTCTCTCTTTAAATGAAGTGAAGTCTCAACTTAA 3173

RESULT 2
CR857633
LOCUS CR857633 3850 bp mRNA linear HTC 12-NOV-2004
DEFINITION Pongo pygmaeus mRNA, cDNA DKFp459H0254 (from clone DKFp459H0254).
ACCESSION CR857633
VERSION CR857633.1 GI:55726270
KEYWORDS HTC.
SOURCE Pongo pygmaeus (orangutan)
ORGANISM Pongo pygmaeus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homidae; Pongo.
1 (bases 1 to 3850)
Koehler, K., Beyer, A., Mewes, H.W., Well, B., Amid, C., Oeanger, A.,
Podo, G., Han, M. and Wiemann, S.
The German cDNA Consortium
Direct Submission
Submitted (12-NOV-2004) MIPS, Ingolstaedter Landstr.1, D-85764
Neuberberg, GERMANY
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
Sequenced by BMFZ (Biomedical Research Center at the
Heinrich-Heine-University, Dueseldorf/Germany) within the cDNA
Sequencing consortium of the German Genome Project. This clone
(DKFp459H0254) is available at the RZPD Deutsches
Resourcenzentrum fuer Genomforschung GmbH in Berlin, Germany.
Please contact RZPD for ordering:
http://www.rzpd.de/cgi-bin/products/cl.cgi?cloneID=DKFp459H0254
Further information about the clone and the sequencing project is
available at http://mips.gsf.de/projects/cdna/.

FEATURES
source location/Qualifiers
1..3850
/organism="Pongo pygmaeus"
/mol_type="mRNA"
/db_xref="taxon:9600"
/clone="DKFp459H0254"
/tissue_type="cortex"
/clone_lib="459 (synonym: pcor1). Vector pSport1_Sfi; host
DH10B; sites SfiIa + SfiIb"
/dev_stage="adult"
/note="zinc finger protein 331 (Homo sapiens)"
1..3850
/gene="DKFp459H0254"
121..1512
/gene="DKFp459H0254"
/codon_start=1
/product="hypothetical protein"
/protein_id="CAH89907.1"
/db_xref="GI:55726271"
/translation="WAQGLVTFADVAIIFPSQEMACLSAQBDLYMDVLENYLVS
LDLESAYENKSLPEKNIHEIIRASRNSDRSKSLGRWVCEGMLEPQSRGKRVNQ
MIINVVKRPATETPTPTTHORHKNSEKCDCKAFSGYOLSHOKIHGKPEYE
CKECKAFRWGQLOTHOKIHGKPEYCKDCGKAFWGSILVIKRIHTGKPEYCK
DCGKAFRGDELTOHQRPHTGKPEYCKDCGKTFRRVYKLQHKRIHSGEKYEKDCG
KRAPICGSILQHKRIHSGEKYECOBGKAFTRVNYLQHKRIHSGEKYEKDCG
AFRWGSSLVKRIHHTGKPEYCKTECGAFNCGYHLTOHERIHTGKPEYCKEKGAF
IYGSLSVHERIHTGKPEYCKTECGKSFSGHQLTOHKTHSGANSYCKEKGACNH
LNHLNEHORIHNS"

ORIGIN
Query Match 91.6%; Score 1845.6; DB 4; Length 3850;
Best Local Similarity 97.7%; Pred. No. 0;
Matches 1905; Conservative 0; Mismatches 39; Indels 6; Gaps 3;

QY 72 ATCCCGAGAGAGAGAGAGATCGTTAAACATCTTA---GTCAGCTCAGCCTCTCGA 128
DB 1 ATCCCGAGAGAGAGAGAGATCGTTAAACATCTTAAGAGTCAAGCTCAGCTCTCGA 60

QY 129 ATT--TGTCTTTCAGTGAAGAAACCCCGAGAACTGATCGAGTCTGCTTAACA 186
DB 61 ATTGTGCTCTTCAGTGAAGAAACCCCGAGAACTGATCGAGTCTGCTTAACA 120

QY 187 ATGGCCAGAGGTTTGTGACGTTCCCGACGTAAGCATGACTTTCTCAGAGAGATGG 246
DB 121 ATGGCCAGAGGTTTGTGACATTCGCGCATGTGACCATAGACTTTCTCAGAGAGATGG 180

QY 247 GCTGTCTGAACCTCTGCTCAAGGAGACTGTACTGGACGTTGATGCTGAGAACTACAGT 306
DB 181 GCTGTCTGAACCTCTGCTCAAGGAGACTGTACTGGACGTTGATGCTGAGAACTACAGT 240

QY 307 AACTGTGCTCAGCTGAGATTTGGAGTCACATATGAAAATTAAGATTTCTTCAAGAAA 366
DB 241 AACTGTGCTCAGCTGAGATTTGGAGTCACATATGAAAATTAAGATTTCTTCAAGAAA 300

QY 367 AACATTCTAATAATGAAGGCTTCCAAAAGAAATTCAGATGAGAAATATCCCTTGGC 426
DB 301 AACATTCTAATAATGAAGGCTTCCAAAAGAAATTCAGATGAGAAATATCCCTTGGC 360

QY 427 CGTAACTGATATGTAAGGTAAGTACGTTGAAAGACCAAGCGCTCCAGAGGAGTATGTC 486
DB 361 CGTAACTGATATGTAAGGTAAGTACGTTGAAAGACCAAGCGCTCCAGAGGAGTATGTC 420

QY 487 AATCAGATGATCATGAATATGTCAAAAGCGCTGCTACTAGAGAGGACCCCTCTGA 546
DB 421 AATCAGATGATCATGAATATGTCAAAAGCGCTGCTACTAGAGAGGACCCCTCTGA 480

QY 547 ACACATCAGAGACATCATGAAGAAATTCCTTGAATGTAAGACTGTGGAGAGGCTTT 606
DB 481 ACACATCAGAGACATCATGAAGAAATTCCTTGAATGTAAGACTGTGGAGAGGCTTT 540

QY 607 AGCTGTGCTATCAACTTAAGTCAACATCAAGAAATTCATCTGTGAGAAAACCTTATGA 666
DB 541 AGCTGTGCTATCAACTTAAGTCAACATCAAGAAATTCATCTGTGAGAAAACCTTATGA 600

QY 667 TGTAAAGATGTAAGAGGCTTCCGTTGGGCAATCGACTTCAACATCAAAAAAT 726
DB 601 TGTAAAGATGTAAGAGGCTTCCGTTGGGCAATCGACTTCAACATCAAAAAAT 660

QY 727 CATATCGGGAGAAAGCCCTCAAGATGTAAAGACTGTGGAGAGGCTTTGATGGGGCTCA 786
DB 661 CATATCGGGAGAAAGCCCTCAAGATGTAAAGACTGTGGAGAGGCTTTGATGGGGCTCA 720

QY 787 AGCTGTGATTCATTAAGAGATTCATCTGTGAAAAACCTATGAATGTAAGACTGT 846
DB 721 AGCTGTGATTCATTAAGAGATTCATCTGTGAAAAACCTATGAATGTAAGACTGT 780

QY 847 GGAAGAGGCTTTCGGCGGTGTGATGAGTCACTGACGACAGAGATTCACACTGGGAG 906
DB 781 GGAAGAGGCTTTCGGCGGTGTGAGTCACTGACGACAGAGATTCACACTGGGAG 840

QY 907 AAAGACTAGAGATGGAAGACTGTGGAGAAAGCTTTAGCCGTTGATTAACCTTATTCAG 966
DB 841 AAAGACTAGAGATGGAAGACTGTGGAGAAAGCTTTAGCCGTTGATTAACCTTATTCAG 900

QY 967 CACAAGAAATTCATTAAGTGGGAGAAAGCTTACGAGTGAAGAACTGTGGAGAGCTTT 1026
DB 901 CACAAGAAATTCATTAAGTGGGAGAAAGCTTACGAGTGAAGAACTGTGGAGAGCTTT 960

QY 1027 ATTTGTGTTCAAGCTCATTCAGATTAAGAAATTCACAGGTGAGAAAGCTTATGAA 1086
DB 961 ATTTGTGTTCAAGCTCATTCAGATTAAGAAATTCACAGGTGAGAAAGCTTATGAA 1020

QY 1087 TGTCAAGATGTTGGAGAGGCTTTACTCGAGTCAATTAACCTTACAGATCAGAAAGTC 1146
DB 1021 TGTCAAGATGTTGGAGAGGCTTTACTCGAGTCAATTAACCTTACAGATCAGAAAGTC 1080

QY 1147 CACACGGTGAAGAGCTTCAAGATGTAAGAGTGTGGAGAGGCTTTGCTGGGGTTCG 1206
DB 1081 CACACGGTGAAGAGCTTCAAGATGTAAGAGTGTGGAGAGGCTTTGCTGGGGTTCG 1140

```

QY 1207 AGCTCGTTAAGCAGAGAGATACATACGGGCGAGAAACCGGTACAGTGCACAGATAT 1266
DB 1141 AGCTCGTTAAGCAGAGAGATACATACGGGCGAGAAACCGGTACAGTGCACAGATAT 1200
QY 1267 GGGAAAGCCTTCAATTGTGGCTATCACTCACTCAGCAGAGAAATCCACACAGCGAA 1326
DB 1201 GGGAAAGCCTTCAATTGTGGCTATCACTCACTCAGCAGAGAAATCCACACAGCGAA 1260
QY 1327 ACCCGGTAAATTAAGAGAGTGGGAAAGCGCTTCAATTAGATCCAGCGCTCGGAA 1386
DB 1261 ACCCGGTAAATTAAGAGAGTGGGAAAGCGCTTCAATTAGATCCAGCGCTCGGAA 1320
QY 1387 CATGAGAGATTCATACCGGGGTGAACCTTAATGGGTATACAGATGTGGAAAGACTTT 1446
DB 1321 CATGAGAGATTCATACCGGGGTGAACCTTAATGGGTATACAGATGTGGAAAGACTTT 1380
QY 1447 AGTCACGGCCATCAGCTTACACAACTCAGAAACCGCAGTGGGGCGAAATCTTACGAA 1506
DB 1381 AGTCATGGCCATCAGCTTACACAACTCAGAAACCGCAGTGGGGCGAAATCTTACGAA 1440
QY 1507 TGTAAAGAGTGGGGGAAAGGCAATGTAACCACTTAAACATCTCCGAGAACATCAGAGATC 1566
DB 1441 TGTAAAGAGTGGGGGAAAGGCGGTGTAAACCACTTAAACATCTCCGAGAACATCAGAGATC 1500
QY 1567 CACAAAGTTGAAGAGCCTTTTGAACGAGTAGCCGCTGTATCTATGTTTGGCTTTC 1626
DB 1501 CACAAAGTTGAAGAGCCTTTTGAACGAGTAGCCGCTGTATCTATGTTTGGCTTTC 1560
QY 1627 CACAGTTTGTATCTGCACTCAACCTCAGTTCAAAATTTAAATGAAAAATTCAGAAA 1686
DB 1561 CACAGTTTGTATCTGCACTCAACCTCAGTTCAAAATTTAAATGAAAAATTCAGAAA 1620
QY 1687 TAAAGAAATT- TAAGTCTCAAAATGTTGCGCTTGAAGAGGTATGAATCTCTCGC 1745
DB 1621 CAAAGAAATTGTATGTTCTCAATGTTGCGCTTGAAGAGGTATGAATCTCTCGC 1680
QY 1746 TGTCCGCTCAAGCCGCGGGAGTGAATCATCCCTTGTGCAGACATCCACGCTGT 1805
DB 1681 TGTCCGCTCAAGCCGCGGGAGTGAATCATCCCTTGTGTCCAGACATCCACGCTGT 1740
QY 1806 ATACGCCACCCACCTGTGTATGATCTTAAGTCCGTCTGTGTATGATCAATCAATCCC 1865
DB 1741 ACACGCTACCCACCTGTGTATGATCTTAAGTCCGTCTGTGTATGATCAATCAATCCC 1800
QY 1866 AGATATCAGTCCGTGTGTGTGCAAGTGTCTCTTGTCTTACAGTGGCCCGAGAGAC 1925
DB 1801 AGATATCAGTGTGTGTGTGCTCAAGTGTCTTCTTGTCTTACAGTGGCCCGAGAGAC 1860
QY 1926 AGAGATAGTATGCTGTGTATGATGATGCAAGAGAGACCAAGTGTCTTCTTCTT 1985
DB 1861 AGAGATAGTATGCTGTGTATGATGATGCAAGAGAGACCAAGTGTCTTCTTCTT 1920
QY 1986 AAATGAAAAGGTGAAAGTTCTCAACTTAA 2015
DB 1921 AAATGAAAAGGTGAAAGTTCTCAACTTAA 1950

RESULT 3
DQ045794 1383 bp DNA linear GSS 02-JUN-2005
LOCUS Homo sapiens ZNF463 gene, VIRUTAL TRANSCRIPT, partial sequence,
DEFINITION genomic survey sequence.
ACCESSION DQ045794
VERSION DQ045794.1 GI:66897009
KEYWORDS GSS.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominoidea; Homo.
1 (bases 1 to 1383)
Nietzen, R., Bustamante, C., Clark, A.G., Glanowski, S., Sackton, T.B.,
Hubisz, M.J., Fiedel-Alon, A., Tanenbaum, D.M., Civejlo, D.,

```

```

TITLE
JOURNAL White, T.J., Shinsky, J.J., Adams, M.D. and Cargill, M.
PUBMED A Scan for Positively Selected Genes in the Genomes of Humans and
AUTHORS (et) PLoS Biol. 3 (6), E170 (2005)
REFERENCE 15869325
2 (bases 1 to 1383)
Nietzen, R., Bustamante, C., Clark, A.G., Glanowski, S., Sackton, T.B.,
Hubisz, M.J., Fiedel-Alon, A., Tanenbaum, D.M., Civejlo, D.,
White, T.J., Shinsky, J.J., Adams, M.D. and Cargill, M.
TITLE Direct Submission
JOURNAL Submitted (05-MAY-2005) Celera Genomics, 45 West Gude Drive,
COMMENT Rockville, MD 20850, USA
This sequence was made by sequencing genomic exons and ordering
them based on alignment. Translation starts at the beginning of
alignment.
FEATURES
source
location/Qualifiers
1..1383
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="19"
<1..>1383
/gene="ZNF463"
/locus_tag="HC11559"
gene
ORIGIN
Query Match 68.6%; Score 1383; DB 11; Length 1383;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1383; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 196 GGTGGTGTGACGTTGCGCGACGTACCATAGACTTTTCTCAGAGAGAGTGGCGTGTCTG 255
DB 1 GGTGGTGTGACGTTGCGCGACGTACCATAGACTTTTCTCAGAGAGAGTGGCGTGTCTG 60
QY 256 AACTGTGCTCAAGAGGACCTGTATCTGGAGCGTGAATGCTGGAGAAATCACTAATCTTGTGC 315
DB 61 AACTGTGCTCAAGAGGACCTGTATCTGGAGCGTGAATGCTGGAGAAATCACTAATCTTGTGC 120
QY 316 TCACGTGATTTGAGTGCAGCATATGAAAAATAGAGTTTACTTACAGAAAAAAATTCAT 375
DB 121 TCACGTGATTTGAGTGCAGCATATGAAAAATAGAGTTTACTTACAGAAAAAAATTCAT 180
QY 376 GAAATAGAGGCTTCCAAAAGAAATTCAGATAGAGAAATCCCTTGGCGGTAACTGG 435
DB 181 GAAATAGAGGCTTCCAAAAGAAATTCAGATAGAGAAATCCCTTGGCGGTAACTGG 240
QY 436 ATATGTGAAGGTATGCTTGAAGAACCAAGCGCTCCAGAGGAGGTATGTCAATCAGATG 495
DB 241 ATATGTGAAGGTATGCTTGAAGAACCAAGCGCTCCAGAGGAGGTATGTCAATCAGATG 300
QY 496 ATCATCAATTAATGTCAAAAAGGCTGTACTATAGAGAGCAACCTCTAGAAACATCAG 555
DB 301 ATCATCAATTAATGTCAAAAAGGCTGTACTATAGAGAGCAACCTCTAGAAACATCAG 360
QY 556 AGACATCATAGAGAGATTCCTTTGAATGTAAAGACTGTGGAAAGCGCTTTAGTGTGTC 615
DB 361 AGACATCATAGAGAGATTCCTTTGAATGTAAAGACTGTGGAAAGCGCTTTAGTGTGTC 420
QY 616 TATCAACTTATGTCAACATCAGAAAAATCCATCTGTGTGAGAAAACTTATGAATGTAAGA 675
DB 421 TATCAACTTATGTCAACATCAGAAAAATCCATCTGTGTGAGAAAACTTATGAATGTAAGA 480
QY 676 TGTAAAGAGCGCTTCGTTGGGGCAATCAGTTTACTCAATCAAAAATTCATCTGGG 735
DB 481 TGTAAAGAGCGCTTCGTTGGGGCAATCAGTTTACTCAATCAAAAATTCATCTGGG 540
QY 736 GAGAAAGCCTTCAAGATGTAAAGACTGTGTGAAAGCGCTTTTCAATGGGCTCAAGCTCGTT 795
DB 541 GAGAAAGCCTTCAAGATGTAAAGACTGTGTGAAAGCGCTTTTCAATGGGCTCAAGCTCGTT 600
QY 796 ATTCAATAGAGATTTATCTGTGTGAAAAACCTTATGAATGTAAAGACTGTGAAAGGCC 855
DB 601 ATTCAATAGAGATTTATCTGTGTGAAAAACCTTATGAATGTAAAGACTGTGAAAGGCC 660

```

```

QY 856 TTTCCGCGGTGATGAGTCACTGACGACCAAGATTCCACACTGGGAGAAAGACTAC 915
Db 661 TTTCCGCGGTGATGAGTCACTGACGACCAAGATTCCACACTGGGAGAAAGACTAC 720
QY 916 GAATGCAAGACTGTGGGAGACCTTTAGCGGTGTATTAACCTTATTGAGACAAGAGA 975
Db 721 GAATGCAAGACTGTGGGAGACCTTTAGCGGTGTATTAACCTTATTGAGACAAGAGA 780
QY 976 ATTCAATGAGGAGAGAGCTTACAGATGTGTAAAGACTGTGGAGAGCTTTATTGTGT 1035
Db 761 ATTCAATGAGGAGAGAGCTTACAGATGTGTAAAGACTGTGGAGAGCTTTATTGTGT 840
QY 1036 TCAACCTTCATTCAGATTAAGAAATTCACACAGGTGAGAAAACCTTAATGATGCAAGAA 1095
Db 841 TCAACCTTCATTCAGATTAAGAAATTCACACAGGTGAGAAAACCTTAATGATGCAAGAA 900
QY 1096 TGTGGGAGAGGCTTTACTGAGTCAATTACCTTACTCAGCATCAGAAAGATCCACACCGGT 1155
Db 901 TGTGGGAGAGGCTTTACTGAGTCAATTACCTTACTCAGCATCAGAAAGATCCACACCGGT 960
QY 1156 GAGAAAGCTCAGAAATGTAAGAGTGTGGAAAGGCTTTGCTGGGCTTCGAGCCTCGTT 1215
Db 961 GAGAAAGCTCAGAAATGTAAGAGTGTGGAAAGGCTTTGCTGGGCTTCGAGCCTCGTT 1020
QY 1216 AAGCAGAGAGATCATTCGCGGCGAGAAAGCCTTACAGTCAAGATGTGGAAAGGCC 1275
Db 1021 AAGCAGAGAGATCATTCGCGGCGAGAAAGCCTTACAGTCAAGATGTGGAAAGGCC 1080
QY 1276 TTCAATTTGTGCTATCACCTCACTCAGACGAGAGATCCACAGGCGGAAACCCCGTAT 1335
Db 1081 TTCAATTTGTGCTATCACCTCACTCAGACGAGAGATCCACAGGCGGAAACCCCGTAT 1140
QY 1336 AAATGTAAAGAGTGTGGAGAGCTTTTATGATCAGAGCCTGTGAAACATGAGAGA 1395
Db 1141 AAATGTAAAGAGTGTGGAGAGCTTTTATGATGAGAGCCTGTGAAACATGAGAGA 1200
QY 1396 ATTCAATCCGCGGTGAAACCTTATGAGTGTACAGAAATGTGGAAAGCTTTAGTCAAGGC 1455
Db 1201 ATTCAATCCGCGGTGAAACCTTATGAGTGTGTACAGAAATGTGGAAAGCTTTAGTCAAGGC 1260
QY 1456 CATCAGCTTACACATCATGAGAAACGCAAGTGGGCGGAATCCTTAAGAAATGTAAGAG 1515
Db 1261 CATCAGCTTACACATCATGAGAAACGCAAGTGGGCGGAATCCTTAAGAAATGTAAGAG 1320
QY 1516 TCGCGGAGAGCATGTAAACCTTAACCATCTCCGAGACATCAGAGGATCCACAACAGT 1575
Db 1321 TCGCGGAGAGCATGTAAACCTTAACCATCTCCGAGACATCAGAGGATCCACAACAGT 1380
QY 1576 TGA 1578
Db 1381 TGA 1383

```

```

RESULT 4
LOCUS BU944114 935 bp mRNA linear EST 18-OCT-2002
DEFINITION AGENCOURT 10544811 NIH MGC 107 Homo sapiens cDNA clone
IMAGE:6728338 5', mRNA sequence.
ACCESSION BU944114
VERSION BU944114.1 GI:24132933
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homidae; Homo.
REFERENCE 1 (bases 1 to 935)
AUTHORS NIH-MGC http://mgi.mgi.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov

```

Tissue Procurement: ATCC
 cDNA Library Preparation: Rubin Laboratory
 DNA Sequencing by: The I.M.A.G.E. Consortium (LNL)
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
 http://image.lnl.gov
 Plate: LNCM3049 row: m column: 09
 High quality sequence stop: 669.
 Location/Qualifiers
 1..935
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:6728338"
 /tissue_type="adenocarcinoma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NIH MGC 107"
 /note="Organ: breast; Vector: pOTB7; Site 1: ECORI;
 Site 2: XhoI; cDNA made by oligo-dT priming.
 Directionally cloned into EcoRI/XhoI sites using the
 following 5' adaptor: GGCACAG(G). Library constructed by
 Ling Hong in the laboratory of Gerald M. Rubin (University
 of California, Berkeley) using ZAP-cDNA synthesis kit
 (Stratagene) and Superscript II RT (Life Technologies).
 Note: this is a NIH-MGC Library."

ORIGIN

Query Match 43.0%; Score 867; DB 5; Length 935;
 Best Local Similarity 99.0%; Pred. No. 5; 6e-236;
 Matches 903; Conservative 0; Mismatches 6; Indels 3; Gaps 3;

```

QY 187 ATGCCAGGGGTTGTGACGTTGCGGACGTAAGCATTAAGCTTTCTCAGAGAGAGTGG 246
Db 1 ATGCCAGGGGTTGTGACGTTGCGGACGTAAGCATTAAGCTTTCTCAGAGAGAGTGG 60
QY 247 GCGTGTGAACTCTGCTCAGAGGACCTGTACTGGAGCGTATGCTGTGAGAACTACACT 306
Db 61 GCGTGTGAACTCTGCTCAGAGGACCTGTACTGGAGCGTATGCTGTGAGAACTACACT 120
QY 307 AACTGTGCTCAGCTGATTTGGAAGTCAAGATATGAAATTAAGATTACTACAGAAAAA 366
Db 121 AACTGTGCTCAGCTGATTTGGAAGTCAAGATATGAAATTAAGATTACTACAGAAAAA 180
QY 367 AACATTCAATGAATTAAGGCTTCCAAAAGAAATTCAGATAGAAAGTAATCCCTTGGC 426
Db 181 AACATTCAATGAATTAAGGCTTCCAAAAGAAATTCAGATAGAAAGTAATCCCTTGGC 240
QY 427 CGTAACTGGATATGTAAGAGTACGTTGAAAGACCAAGCGCTCCAGAGGAGGTATGTC 486
Db 241 CGTAACTGGATATGTAAGAGTACGTTGAAAGACCAAGCGCTCCAGAGGAGGTATGTC 300
QY 487 AATCAGATGATCATCAATTAATGTCAAAAGGCTGCTACTAGAGAGGACCCCTCTAGA 546
Db 301 AATCAGATGATCATCAATTAATGTCAAAAGGCTGCTACTAGAGAGGACCCCTCTAGA 360
QY 547 ACACATCAGAGACATCATTAAGAGAAATTCCTTGAATGTAAAGACTGTGGAGAGGCTTT 606
Db 361 ACACATCAGAGACATCATTAAGAGAAATTCCTTGAATGTAAAGACTGTGGAGAGGCTTT 420
QY 607 AGTGTGCTATCAACTTAATGTCACATCAAGAAATTCATCTGTGAGAGAAACCTTATGAA 666
Db 421 AGTGTGCTATCAACTTAATGTCACATCAAGAAATTCATCTGTGAGAGAAACCTTATGAA 480
QY 667 TGTAAAGATGTAAAGAGGCTTCCGTTGGGCGCATCAGCTTACTCAACATCAAAAAAT 726
Db 481 TGTAAAGATGTAAAGAGGCTTCCGTTGGGCGCATCAGCTTACTCAACATCAAAAAAT 540
QY 727 CATTAATGGGAGAGGCTTACAGATGTAAAGACTGTGGAGAGGCTTTTCGATGGGCTCA 786
Db 541 CATTAATGGGAGAGGCTTACAGATGTAAAGACTGTGGAGAGGCTTTTCGATGGGCTCA 600
QY 787 AGCTCGTTATTCATTAAGAGATTCATCTGTGTAAAGAAACCTTATGAATGTAAAGACTGT 846

```

```

|||||
Db 601 ACCCTGTTATTCATGAAGAGATTCACTGCTGGAAGAAACCTTATGAATGAATGAAGACTGT 660
QY 847 GGAAGAGGCTTTTCGGGTGTGATGAGCTCACTCAGACACAGAGATTCCACACTGG-6GA 905
Db 661 GGAAGAGGCTTTTCGGGTGTGATGAGCTCACTCAGACACAGAGATTCCACACTGGAG-720
QY 906 GAAAGACTAGAGATGCAAGAGACTGTGGAGAGACCTTTAGCCGTGTATTAATTAATTCA 965
Db 721 GAAAGACTAGAGATGCAAGAGACTGTGGAGAGACCTTTAGCCGTGTATTAATTAATTCA 780
QY 966 GCAAGAGAGATTCATATGCTGGGAGAGAGCTTTACGAGTGTAAAGACTGTGGAGAGC-1T 1024
Db 781 GCAAGAGAGATTCATATGCTGGGAGAGAGCTTTACGAGTGTAAAGACTGTGGAGAGCCTTT 840
QY 1025 TTATTTGTGTTCAACCTCATTCAGCATTAAGAAATTCACACAGGT-GAGAAACCTTAT 1083
Db 841 TTATTTGTGTTCAACCTCATTCAGCATTAAGAAATTCACACAGGTGAGAAACCTTAT 900
QY 1084 GAATGTCAAGAA 1095
Db 901 GAATGTCAAGAA 912

RESULT 5
LOCUS BM014399 839 bp mRNA linear EST 30-OCT-2001
DEFINITION 603640139F1 NIH_MGC_87 Homo sapiens cDNA clone IMAGE:5416496 5',
            mRNA sequence.
ACCESSION BM014399
VERSION BM014399.1 GI:16528753
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Homiidae; Homo.
REFERENCE 1 (bases 1 to 839)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Straubeberg, Ph.D.
            Email: cgabbs-r@mail.nih.gov
            Tissue Procurement: DCTD/DRP
            cDNA Library Preparation: Life Technologies, Inc.
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LNL at:
            http://image.llnl.gov
            Plate: L1AM12062 row: m column: 09
            High quality sequence stop: 818.
FEATURES
    source
        1..839
            location/Qualifiers
                /organism="Homo sapiens"
                /mol_type="mRNA"
                /db_xref="taxon:9606"
                /clone="IMAGE:5416496"
                /tissue_type="mammary adenocarcinoma, cell line"
                /lab_host="DH10B (phage-resistant)"
                /clone_lib="NIH_MGC_87"
                /note="Organ: breast; Vector: PCMV-SPORT6; Site 1: NotI;
                Site 2: SalI; Cloned unidirectionally; oligo-dT primed.
                Average insert size 1.383 kb. Library enriched for
                full-length clones and constructed by Life Technologies.
                Note: this is a NIH_MGC Library."
ORIGIN
Query Match 39.8%; Score 802.2; DB 3; Length 839;
Best local Similarity 98.8%; Pred. No. 1,8e-217;
Matches 829; Conservative 0; Mismatches 8; Indels 2; Gaps 2;
QY 803 AGAGATTCTACTGTGAAAAACCTTATGATGTAAAGACTGTGAAAGGCTTTTCGGC 862

```

```

|||||
Db 1 AGAGATTCTACTGTGAAAAACCTTATGATGTAAAGACTGTGAAAGGCTTTTCGGC 60
QY 863 GTGTGATGAGCTCACTCAGACACAGAGATTCCACACTGGGAGAGAAAGACTAGATGCA 922
Db 61 GTGTGATGAGCTCACTCAGACACAGAGATTCCACACTGGGAGAGAAAGACTAGATGCA 120
QY 922 AAGACTGTGGAGAGACCTTTAGCCGTGTATTAATTAATTCAAGACAAAGAGATTGATA 982
Db 121 AAGACTGTGGAGAGACCTTTAGCCGTGTATTAATTAATTCAAGACAAAGAGATTGATA 180
QY 983 GTGGGAGAGAGCTTTACAGAGTGTAAAGACTGTGGAGAGCCTTTATTTGTGTCAAGCC 1042
Db 181 GTGGGAGAGAGCTTTACAGAGTGTAAAGACTGTGGAGAGCCTTTATTTGTGTCAAGCC 240
QY 1043 TCATTTCAGATTAAGAAATTCACACAGGTGAGAGAAACCTTATGAATGTCAAGATGTGGA 1102
Db 241 TCATTTCAGATTAAGAAATTCACACAGGTGAGAGAAACCTTATGAATGTCAAGATGTGGA 300
QY 1103 AGGCTTTTACTGAGTCAATTACTTACTCAGCATCAGAGATCCACACCGGTGAGAGC 1162
Db 301 AGGCTTTTACTGAGTCAATTACTTACTCAGCATCAGAGATCCACACCGGTGAGAGC 360
QY 1163 CTCAGGAATGTAAAGAGTGTGGAGAGCCTTTTCGCTGGGGTTCGAGCCTGTTAGCAGG 1222
Db 361 CTCAGGAATGTAAAGAGTGTGGAGAGCCTTTTCGCTGGGGTTCGAGCCTGTTAGCAGG 420
QY 1223 AAGAGATACATCGGGGAGAGAGCCTTACAGTGCACAGATGTGGAGAGCCTTCAATT 1282
Db 421 AAGAGATACATCGGGGAGAGAGCCTTACAGTGCACAGATGTGGAGAGCCTTCAATT 480
QY 1283 GTGGCTATCATCTCACTCAGACAGAGAGATCCACACAGGCGAAACCCGTATTAATGTA 1342
Db 481 GTGGCTATCATCTCACTCAGACAGAGAGATCCACACAGGCGAAACCCGTATTAATGTA 540
QY 1343 AGGAGTGTGGAGAGGCTTTCATTATGATTCGAGCTCGTGAAGCATGAGAAATTCATA 1402
Db 541 AGGAGTGTGGAGAGGCTTTCATTATGATTCGAGCTCGTGAAGCATGAGAAATTCATA 600
QY 1403 CCGGGGTGAAGACCTTATGAGGTGTACAGAAATGTGGAGAGCTTTAGTCAAGCATCAGC 1462
Db 601 CCGGGGTGAAGACCTTATGAGGTGTACAGAAATGTGGAGAGCTTTAGTCAAGCATCAGC 660
QY 1463 TTACACACATCATGAGAAAGCAGCAGTGGGCGAAATCTTACAGATGTAAAGAGTGC-6G3 1521
Db 661 TTACACACATCATGAGAAAGCAGCAGTGGGCGAAATCTTACAGATGTAAAGAGTGC-720
QY 1522 AAGGCAATGTAACCACTTAACCATCTCCGAGAAACATCAGAGATCCAC-AAAGTTGAG 1580
Db 721 AAGGCAATGTAACCACTTAACCATCTCCGAGAAACATCAGAGATCCACAAACGTTGAAAG 780
QY 1581 AGCCTTTGAAGCAGATACCCGCTGTATCTATGTTTCGCTTTCACAGTTTGTAC 1639
Db 781 AGCCTTTGAAGCAGATACCCGCTGTATCTATGTTTCGCTTTCACAGATGTTAC 839

RESULT 6
CT001323 909 bp mRNA linear EST 28-JUN-2005
LOCUS CT001323 R2PD no.9016 Homo sapiens cDNA clone R2PDp9016A1714 5',
DEFINITION mRNA sequence.
ACCESSION CT001323
VERSION CT001323.1 GI:68295206
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Homiidae; Homo.
REFERENCE 1 (bases 1 to 909)
AUTHORS Heil,O., Ebert,L., Hennig,S., Henze,S., Radelof,U., Schneider,D.
            and Korn,B.
TITLE Human T-Lymphocytes library

```

JOURNAL
COMMENT

Unpublished (2005)
Contract: Inge Ariart
RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
Heubnerweg 6, D-14059 Berlin, Germany
Email: www.rzpd.de
RZPD; RZPDP9016A1714.
RZPDLIB; (Human T-Lymphocytes) RZPD LIB No. 9016
http://www.rzpd.de/cgi-bin/products/set.cgi?libno=9016 Contact:
Inge Ariart
RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
Heubnerweg 6, D-14059 Berlin, Germany
Tel: +49 30 32639 100
Fax: +49 30 32639 111
www.rzpd.de
This clone is available from RZPD;
http://www.rzpd.de/cgi-bin/products/cl.cgi?cloneid=RZPDP9016A1714
contact RZPD (product-support@rzpd.de) for further information.
Primer name: qe3 4, Primer sequence: CGGATACCAATTCACACAG.

FEATURES

source

1..909
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="caxon:9606"
/clone="RZPDP9016A1714"
/tissue_type="T-Lymphocytes"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="RZPD no.9016"
/note="Vector: pQE80LSN_cloned; Site_1: SalI; Site_2:
NotI; vector:
http://www.rzpd.de/info/vectors/pQE80LSN_cloned.pic.shtml
; 1st strand cDNA was prepared from mRNA obtained from
human T-Lymphocytes with a NotI - 0150 (dt) primer [5'
GACTAGTTCATGATCGCAGCGCCGCTTTTCTTTTCTTTT 3']
Double-stranded cDNA was ligated to SalI adaptors,
digested with NotI and cloned into the NotI and SalI sites
of the pQE80LSN_cloned vector"

ORIGIN

Query Match 37.1%; Score 748; DB 7; Length 909;
Best Local Similarity 98.8%; Pred. No. 5.9e-202;
Matches 770; Conservative 0; Mismatches 7; Indels 2; Gaps 2;

47 CAGGCCAGCATCTTCAAGAAAAAGCATCCCGAGAGAGAGACGAATGTTAAACATCTT 106
131 CAGGCGACATCTTCAAGAAAAAGCATCCCGAGAGAGAGACGAATGTTAAACATCTT 190
107 AGGTGAGCTAGCTCTCGGAATTGTTCTTCTTCAGTGGAACCCGAGAGAACTGATC 166
191 AGGTGAGCTAGCTCTCGGAATTGTTCTTCTTCAGTGGAACCCGAGAGAACTGATC 250
167 AGTCTTCACTTCAAAACAAATGCGCCAGAGGTTGGTACGTTCCGCGACGTAGCCATAG 226
251 AGTCTTCACTTCAAAACAAATGCGCCAGAGGTTGGTACGTTCCGCGACGTAGCCATAG 310
227 ACTTTTCTCAGAGAGAGTGGGCTGTGAACTGTGCTTCAAGAGGACCTGTACTGGAACG 286
311 ACTTTTCTCAGAGAGAGTGGGCTGTGAACTGTGCTTCAAGAGGACCTGTACTGGAACG 370
287 TGATGCTGGAAGAACTACAGTAACTTGCTGCTCACTGGAATTTGGAGTCACAGATTA 346
371 TGATGCTGGAAGAACTACAGTAACTTGCTGCTCACTGGAATTTGGAGTCACAGATTA 430
347 AGAGTTTACCTACAGAAAAAAACATTCAATGAATTAAGGCTTCCAAAAGAAATTCAGATA 406
431 AGAGTTTACCTACAGAAAAAAACATTCAATGAATTAAGGCTTCCAAAAGAAATTCAGATA 490
407 GAAGAAGTAATCCCTTGGCCCTACCTGATATGTGAAGAGTCCGTTGAAGACCAACAGC 466
491 GAAGAAGTAATCCCTTGGCCCTACCTGATATGTGAAGAGTCCGTTGAAGACCAACAGC 550
467 GCTCAGAGAGAGTATGTCATGATGATCATCAATTAATGTCAAAAGGCTGTACTACTA 526

Db 551 GCTCAGAGAGAGTATGTCATGATGATCATCAATTAATGTCAAAAGGCTGTACTACTA 610
Qy 527 GAGAAAGCAACCCCTCTTGAACACATTCAGACATCATTAAGAGAAATTCCTTGAATGA 586
Db 611 GAGAAGGCAACCCCTCTTGAACACATTCAGACATCATTAAGAGAAATTCCTTGAATGA 670
Qy 587 AGGACTGTGGGAAGGCTTGTGCTGCTATCACTTAAGTCAACATCAGAAAAATCCATA 646
Db 671 AGGACTGTGGGAAGGCTTGTGCTGCTATCACTTAAGTCAACATCAGAAAAATCCATA 730
Qy 647 CTGTGAGAAACCTTATGATATGTAAGAATGTAAAGAGCCCTTCGTTGGGCAATCAGC 706
Db 731 CTGTGAGAAACCTTATGATATGTAAGAATGTAAAGAGCCCTTCGTTGGGCAATCAGC 790
Qy 707 TTACTCAACATCAAAAAATTCATACCTGGGAGAGAGCCCTTACGATGTAAAGAC-TGTGGG 765
Db 791 TTACTCAACATCAAAAAATTCATACCTGGGAGAGAGCCCTTACGATGTAAAGACCTTGGG 850
Qy 766 AAGCTTTTCATGAGGGGCTCAAGCCTCGTTATTCATTAAGAGATTCATCTGTGTA 824
Db 851 AAGCTTTTCATGAGGGGCTCAAGCCTCGTTATTCATTAAGAGATTCATCTGTGTA 908

RESULT 7
BU176917
LOCUS BU176917 915 bp mRNA linear EST 04-SEP-2002
DEFINITION AGNCOURT_7941001 NIH_MGC_71 Homo sapiens cDNA clone IMAGE:6155028
5', mRNA sequence.
ACCESSION BU176917
VERSION BU176917.1 GI:2269901
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.
1 (bases 1 to 915)
NIH-MGC http://mgi.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgsrbs-remail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNLN at:
http://image.llnl.gov
Plate: LLM13497 row: a column: 13
High quality sequence stop: 539.

FEATURES

source

1..915
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="caxon:9606"
/clone="IMAGE:6155028"
/tissue_type="leiomyosarcoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_71"
/note="Organ: uterus; Vector: pCMV-Sport6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 2.1 kb."

ORIGIN

Query Match 36.6%; Score 738.4; DB 5; Length 915;
Best Local Similarity 98.0%; Pred. No. 3.3e-199;
Matches 790; Conservative 0; Mismatches 11; Indels 5; Gaps 4;

Qy 1071 TGAGAAACCTATGATGTCAAGAAATGTGGAAAGGCTTTTACTGAGTCAATTACCTTAC 1130
Db 1 TGAGAAACCTATGATGTCAAGAAATGTGGAAAGGCTTTTACTGAGTCAATTACCTTAC 60

```

QY 1131 TCAGCATCAGAAATCCACACCGGTGAGAAAGCTCCAGAAATGTAAGAGTGTGGAGGC 1190
DB 61 TCAGCATCAGAAATCCACACCGGTGAGAAAGCTCCAGAAATGTAAGAGTGTGGAGGC 120
QY 1191 CTTTGGCTGGGGTTCGAGCTGTGTAAGACGAGAGAAATACATACCGGGGAGAAAGCCGTA 1250
DB 121 CTTTGGCTGGGGTTCGAGCTGTGTAAGACGAGAGAAATACATACCGGGGAGAAAGCCGTA 180
QY 1251 CAAGTGCACAGAAATGTGGAAAGCCCTTCAATTGTGGCTATCACTCACTCAGACGAGAG 1310
DB 181 CAAGTGCACAGAAATGTGGAAAGCCCTTCAATTGTGGCTATCACTCACTCAGACGAGAG 240
QY 1311 AATCCACACAGCGCAAAACCCGTAATTAATGTAAGAGTGTGGAAAGCCCTTCAATTATG 1370
DB 241 AATCCACACAGCGCAAAACCCGTAATTAATGTAAGAGTGTGGAAAGCCCTTCAATTATG 300
QY 1371 ATCGAGCCCTGTGAAACATGAGAAATTAATACCGGGGTGAAACCCCTATGGGTATACAA 1430
DB 301 ATCGAGCCCTGTGAAACATGAGAAATTAATACCGGGGTGAAACCCCTATGGGTATACAA 360
QY 1431 ATGTGGGAAGAGCTTTAGTCAAGCCCATGAGCTTACACAACATCAGAAAAACGACAGTGG 1490
DB 361 ATGTGGGAAGAGCTTTAGTCAAGCCCATGAGCTTACACAACATCAGAAAAACGACAGTGG 420
QY 1491 GCGCAAAATCTTAAGAAATGTAAGAGTGTGGAAAGCGCATGTAAACCATCTCCG 1550
DB 421 GCGCAAAATCTTAAGAAATGTAAGAGTGTGGAAAGCGCATGTAAACCATCTCCG 480
QY 1551 AAGAACTCAGAGATCCACAAAGTGAAGAGCCCTTTGAAGCAAGCATGACCTAAACCATCTCCG 1610
DB 481 AAGAACTCAGAGATCCACAAAGTGAAGAGCCCTTTGAAGCAAGCATGACCTAAACCATCTCCG 540
QY 1611 CTATGCTTCGCTTTCACAGTTTGTAACTGAGTCAACTGAGTTCAAAAATATATAA 1670
DB 541 CTATGCTTCGCTTTCACAGTTTGTAACTGAGTCAACTGAGTTCAAAAATATATAA 600
QY 1671 TGGAAAAATTCAGAAATTAAGAAATTTTAAAGTCTCAAAATGAGTGT-GCCCTTGTGAGTACG 1729
DB 601 TGGAAAAATTCAGAAATTAAGAAATTTTAAAGTCTCAAAATGAGTGTGAGTGTGAGTACG 660
QY 1730 TGAATGAATCTCTCGCTGTCGCGCTCCAGCCGCGGGGATGTGAATCAT-CCCTTGTGTC 1788
DB 661 TGAATGAATCTCTCGCTGTCGCGCTCCAGCCGCGGGGATGTGAATCATCCCTTGTGTC 720
QY 1789 CAGCAGATCAGAGCTGTATAGCCCAACCCCTGTAGTGAATTAAGCC-GTCTTGTG 1847
DB 721 CAGCAGATCAGAGCTGTATAGCCCAACCCCTGTAGTGAATTAAGCCGCTTGTG 780
QY 1848 GATCA--GATCAACTATCCAGCATC 1871
DB 781 GATCAGGATCAACTATGCCAATC 806

RESULT 8
BUI39183 940 bp mRNA linear EST 13-SEP-2002
LOCUS AGENCOURT 10215200 NIH-MGC 107 Homo sapiens cDNA clone
DEFINITION IMAGE:6569870 5', mRNA sequence.
ACCESSION BUI39183
VERSION BUI39183.1 GI:22849624
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
            Homo sapiens (human)
            Mammalia; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominoidea; Homo.
            1 (bases 1 to 940)
REFERENCE NIH-MGC http://mgi.nci.nih.gov/.
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
            Email: cgab@r-f@mail.nih.gov
            Tissue Procurement: ATCC

```

```

FEATURES
source
1..940
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6569870"
/tissue_type="adenocarcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_id="NIH-MGC 107"
/note="Organ: breast; Vector: pOTB; Site_1: EcoRI;
Site_2: XhoI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCACGAG(G). Library constructed by
ling Hong in the laboratory of Gerald M. Rubin (University
of California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH-MGC Library."

ORIGIN
Query Match 36.3%; Score 730.6; DB 5; Length 940;
Best Local Similarity 95.9%; Pred. No. 5,7e-197;
Matches 793; Conservative 0; Mismatches 29; Indels 5; Gaps 4;

QY 211 GCCGACGTAGCCCATAGACTTTTCTCAGAGAGAGTGGCCCTGTCTGAACCTGTCTCAGAG 270
DB 1 GCCGACGTAGCCCATAGACTTTTCTCAGAGAGAGTGGCCCTGTCTGAACCTGTCTCAGAG 60
QY 271 GACCTGTACTGGGACGTGATGCTGAGAACTACATTAAGTGTCTCACTGATTTGAG 330
DB 61 GACCTGTACTGGGACGTGATGCTGAGAACTACATTAAGTGTCTCACTGATTTGAG 120
QY 331 TCAGCATCAGAAATTAAGAGTTTCTCAGAAAAAAACATTCATGAATAAGGCTTCC 390
DB 121 TCAGCATCAGAAATTAAGAGTTTCTCAGAAAAAAACATTCATGAATAAGGCTTCC 180
QY 391 AAAAGAAATTCAGATAGAAAGTAATCCCTTGGCCGCTGATGTGAAGTACG 450
DB 181 AAAAGAAATTCAGATAGAAAGTAATCCCTTGGCCGCTGATGTGAAGTACG 240
QY 451 CTTGAAAGCCACAGCGCTCCAGAGAGGTATGTCAATCAGATCATTAATATGTC 510
DB 241 CTTGAAAGCCACAGCGCTCCAGAGAGGTATGTCAATCAGATCATTAATATGTC 300
QY 511 AAAAGCCCTGCTACTAGAGAAAGCAACCCCTCTAGAAACATCAGAGACATCATAGAG 570
DB 301 AAAAGCCCTGCTACTAGAGAAAGCAACCCCTCTAGAAACATCAGAGACATCATAGAG 360
QY 571 AATTCCTTTGAATGTAGAGACTGTGGAAAGGCTTTAGTGTGGCTATCACTTACTGCA 630
DB 361 AATTCCTTTGAATGTAGAGACTGTGGAAAGGCTTTAGTGTGGCTATCACTTACTGCA 420
QY 631 CATCAGAAAAATCATCTGTGTGAGAAACCTTATGAATGTAAGAAATGTAAGAGGCTTCC 690
DB 421 CATCAGAAAAATCATCTGTGTGAGAAACCTTATGAATGTAAGAAATGTAAGAGGCTTCC 480
QY 691 CGTTGGGGCAATCAGCTTACTCAACATCAAAAAATTCATCTGAGGAGAAAGCCCTACGA 750
DB 481 CGTTGGGGCAATCAGCTTACTCAACATCAAAAAATTCATCTGAGGAGAAAGCCCTACGA 540
QY 751 TGTAAAGACTGTGGAGAGCTTTTGTAGTGGGCTCAGAGCTGTATTCATTAAGAGATT 810
DB 541 TGTAAAGACTGTGGAGAGCTTTTGTAGTGGGCTCAGAGCTGTATTCATTAAGAGATT 600
QY 811 CATCTGTGAAAAAACCTTATGAATGTAAAGCTGTGAGAAAGGCTTTTGGGGTGTAT 870

```


Db 601 CATACGTGAAAAACCTATGAATGAATAAGACTGTGAAAAAGCTTTGGGGTGTGAT 660

QY 871 GAGCTCATCTGACACCAAGATTCACACTGGGGGAAAAAGACTAGATGCAAAAGCTGT 930

Db 661 GACCTGACTCAG-ACCAGAGA-TCCCACTGGGGGAAAAAGACTAGAAATGCCACACTGT 718

QY 931 GGGAAAGACTTTAAGCCGTGTGTAT-AACTTATTCAGACAAAGAAATTCATAGTGGG- 988

Db 719 GGGAAAGACTTTAAGCCGGGTGTATAAACTTATTCGACAAAGAAATTCATAGTGGG 778

QY 989 -AGAACCTTACGAGTGTAAAGACTGTGGGAAGGCTTTTATTTGTG 1034

Db 779 AAAAACCTTCCGAGTGTAAAGACTTGGGGGAAAGGCTTTTATTTGG 825

RESULT 9
AM376896/c 756 bp mRNA linear EST 04-FEB-2000
LOCUS IL3-CT0219-271099-022-E04 CT0219 Homo sapiens cDNA, mRNA sequence.
ACCESSION AM376896.1 GI:6881559
VERSION AM376896.1 GI:6881559
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.
1 (bases 1 to 756)
HCGP <http://www.ludwig.org.br/ORESTES>.
The FAPESP/LICR Human Cancer Genome Project
Unpublished (1999)
Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(<http://www.ludwig.org.br/scripts/gethtml2.pl?cl=IL3&e2=IL3-CT0219-271099-022-E04&e3=1999-10-27&e4=1>)
Seq primer: puc 18 forward
High quality sequence score: 3
High quality sequence ebp: 669.
Location/Qualifiers
1..756
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_lib="CT0219"
/note="Organ: colon; Vector: puc18; Site 1: Sma1; Site 2:
Sma1. A mini-library was made by cloning products derived
from ORESTES PCR (U.S. Letters Patent application No.
196,716 - Ludwig Institute for Cancer Research) profiles
into the pUC 18 vector. Reverse transcription of tissue
mRNA and cDNA amplification were performed under low
stringency conditions."

ORIGIN
Query Match 35.6%; Score 718.2; DB 1; Length 756;
Best Local Similarity 98.0%; Pred. No. 1.9e-193;
Matches 738; Conservative 0; Mismatches 13; Indels 2; Gaps 1;

QY 703 CAGCTTACCAACATCAAAAATTCATCTGGGAGAGCCCTAGAAATGTAAGACTGT 762

Db 755 CAGCTTTCACACATCAAAA--TTCATATGGGAAGACCTACGAATGTAAGACTGT 698

QY 763 GGGAAAGCTTTTGAATGGGGCTCAAGCCCTGTTATTCATTAAGAGATTCTAGTGTGA 822

Db 697 GGGAAAGCTTTTGAATGGGGCTCAAGCCCTGTTATTCATTAAGAGATTCTAGTGTGA 638

QY 823 AAACCTATGAATGTAAGACTGTGAAAAGGCTTTGGCGGTGTGATGACTCACTAG 882

Db 637 AAACCTATGAATGTAAGACTGTGAAAAGGCTTTGGCGGTGTGATGACTCACTAG 578

QY 883 CACAGAGATTCCACACTGGGGAAGAAAGACTAGAAATGCAAAAGACTGTGGAAAGCTTT 942

Db 577 CACAGAGATTCCACACTGGGGAAGAAAGACTAGAAATGCAAAAGACTGTGGAAAGCTTT 518

QY 943 AGCGTGTGTATTAACCTTATTCAGACAAAGAAATTCATAGTGGGGAAGACTTACAG 1002

Db 517 AGCGTGTGTATTAACCTTATTCAGACAAAGAAATTCATAGTGGGGAAGACTTACAG 458

QY 1003 TGTAAAGCTGTGGGAAGGCTTTTATTTGTGGTCAAGCTCATTCAGATTAAGAAAT 1062

Db 457 TGTAAAGCTGTGGGAAGGCTTTTATTTGTGGTCAAGCTCATTCAGATTAAGAAAT 398

QY 1063 CACACAGGTGAGAAACCTATGAATGTCAAGAAATGTGGGAAGGCTTTTACTGACTCAAT 1122

Db 397 CACACAGGTGAGAAACCTATGAATGTCAAGAAATGTGGGAAGGCTTTTACTGACTCAAT 338

QY 1123 TACCTTACTCAGCATCAGAAATTCACACCGGTGAAAGCCTCAGAAATGTAAAGACTGT 1182

Db 337 TACCTTACTCAGCATCAGAAATTCACACCGGTGAAAGCCTCAGAAATGTAAAGACTGT 278

QY 1183 GGGAAAGGCTTTTCGTGGGGGTTCAGACCTCGTTAAGCAAGAGGATTCATACGGGGGAG 1242

Db 277 GGGAAAGGCTTTTCGTGGGGGTTCAGACCTCGTTAAGCAAGAGGATTCATACGGGGGAG 218

QY 1243 AAGCGTCAAGTGCAGAAATGTGGGAAGGCTTCAATTTGTGCTATCAGCTCAGTCAAG 1302

Db 217 AAGCGTCAAGTGCAGAAATGTGGGAAGGCTTCAATTTGTGCTATCAGCTCAGTCAAG 158

QY 1303 CACGAGAAATTCACACACGCGGAAACCCGTATTAATGTAAAGAGTGTGGGAAGCTTTC 1362

Db 157 CACGAGAAATTCACACACGCGGAAACCCGTATTAATGTAAAGAGTGTGGGAAGCTTTC 98

QY 1363 AATTATGATGAGGCTGTGAAAATGAGAAATTCATCCGGGGTGAACCTATGGG 1422

Db 97 AATTATGATGAGGCTGTGAAAATGAGAAATTCATCCGGGGTGAACCTATGGG 38

QY 1423 TGTACAGAAATGTGGGAAGGCTTTAGTCACGAGC 1455

Db 37 TGTACAGAAATGTGGGAAGGCTTTAGTCACGAGC 5

RESULT 10
B0719354 890 bp mRNA linear EST 16-JUL-2002
LOCUS AGENCOURT_8291371_lupeki_sympathetic_trunk Homo sapiens cDNA clone
DEFINITION IMAGE:6194787 5', mRNA sequence.
ACCESSION B0719354
VERSION B0719354.1 GI:21858251
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.
1 (bases 1 to 890)
REFERENCE NIH-MGC <http://mgs.nci.nih.gov/>.
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE Unpublished (1999)
JOURNAL Contact: Robert Strassberg, Ph.D.
COMMENT Email: csagbs-remail.nih.gov
Tissue Procurement: Dr. James R. Lupski
cDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
<http://image.llnl.gov>
Plate: LLM13600 row: j column: 04

High quality sequence stop: 556.
Location/Qualifiers

1. 890
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6194787"
/sex="male"
/tissue_type="sympathetic trunk"
/dev_stage="adult, 16 Yr"
/lab_host="DH10B"
/clone_lib="lupski_sympathetic_trunk"
/note="Vector: PCWV-SPORT6 (Life Technologies); Site_1:
Noti; Site_2: SalI; cDNA made by oligo-dT priming.
Directionally cloned using the following adaptors:
5'-TGACCCACGCGCTCG-3' and
5'-GACTAGTTCTAGATCGGAGCGCGCCCT(15)-3'. Size selected >
1 kb for average insert length 1.9 kb. This is a primary
library, non-amplified. Library constructed by Life
Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor
College of Medicine); available through Life
Technologies."

ORIGIN

Query Match 35.3%; Score 711.6; DB 5; Length 890;
Best Local Similarity 98.2%; Pred. No. 1.5e-191;
Matches 752; Conservative 0; Mismatches 9; Indels 5; Gaps 3;

394 AGGAATTCAGATGAGAAAGTAATCCCTTGCCGTAAGTATGTAAGTACGCTT 453
1 ATGAATTCAGATGAGAAAGTAATCCCTTGCCGTAAGTATGTAAGTACGCTT 60
454 GAAAGCCACAGGCTCCAGAGGAGTATGTCATACATGATGTCATCAATTAATGCA 513
61 GAAAGCCACAGGCTCCAGAGGAGTATGTCATACATGATGTCATCAATTAATGCA 120
514 AGGCTCTACTAGAGAAAGCAACCTCTAGAACATCAGAGACATCATAGAGAAAT 573
121 AGGCTCTACTAGAGAAAGCAACCTCTAGAACATCAGAGACATCATAGAGAAAT 180
574 TCCCTTTGAATGTAAGACCTGTGGAGAGCCTTTAGTCGTGCTATCAACTTACACAT 633
181 TCCCTTTGAATGTAAGACCTGTGGAGAGCCTTTAGTCGTGCTATCAACTTACACAT 240
634 CAGAATTCATCTGTGAGAAACCTTTAGTGAATGTAAGAAATGAAGAGCCTTCCT 693
241 CAGAATTCATCTGTGAGAAACCTTTAGTGAATGTAAGAAATGAAGAGCCTTCCT 300
694 TGGGCAATCAGCTTCTCAACATCAAAAATTCATCTGGGAGAAAGCCTTACGATGT 753
301 TGGGCAATCAGCTTCTCAACATCAAAAATTCATCTGGGAGAAAGCCTTACGATGT 360
754 AAAAGCTGTGGAGAGCCTTTGATGGGCTCAAGCCTGTTATTCATTAAGAGATTCAT 813
361 AAAAGCTGTGGAGAGCCTTTGATGGGCTCAAGCCTGTTATTCATTAAGAGATTCAT 420
814 ACTGTGAAAAACCTTATGTAATGTAAGACTGTGAGAAAGCCTTTGCGCTGTGATGAG 873
421 ACTGTGAAAAACCTTATGTAATGTAAGACTGTGAGAAAGCCTTTGCGCTGTGATGAG 480
874 CTCACCTCAGACAGAGATTCACACTGGGAGAAAGACTACAGAAATGCAAAAGCTGTGG 933
481 CTCACCTCAGACAGAGATTCACACTGGGAGAAAGACTACAGAAATGCAAAAGCTGTGG 540
934 AAGACCTTTAGCCGTGTGTAATACTTATTCAGACAAAGAAATTCATATGGGAGAG 993
541 AAGACCTTTAGCCGTGTGTAATACTTATTCAGACAAAGAAATTCATATGGGAGAG 600
994 CCTTAGAGGTAAAGCTGTGGAGAGCCTTTATTTGTGTTCAAGCCCTCATTCAGCAT 1053
601 CCTTAGAGGTAAAGCTGTGGAGAGCCTTTATTTGTGTTCAAGCCCTCATTCAGCAT 660
1054 AAAAGATTCACACAGGTGAGAAA--CCTTATGATGTCAAGAA--TGTGGAGAGGCTT 1109

Db 661 AAAAGATTCACACAGGTGAGAAACCTTATGATGTCAAGAAATGTTGGGAGAGCCTT 720

Qy 1110 TACTGATCAATTAACCTTACTCA--GCATCAGAGATGCACACCGG 1154
Db 721 TACTGATCAATTAACCTTACTCAAGGCAATCAGAAATGCACACCGG 766

RESULT 11

CX869347/c

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

1 (bases 1 to 731)

NIH-MGC http://mgc.nci.nih.gov/.

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Other ESTs: HESCA.34.G02.g1.A037

Contact: Daniela S. Gerhard, Ph.D.

Office of Cancer Genomics

National Cancer Institute / NIH

Bldg. 31 Rm10A07 Bethesda, MD 20892

Email: cgapbs-remail.nih.gov

Tissue Procurement: BreaGen, Inc.

cDNA Library Preparation: Express Genomics, Inc.

DNA Sequencing Arrayed by: The I.M.A.G.E. Consortium (LLNL).

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: L1M15782 row: m column: 04

Seq primer: M13-FOR (GTAAACGACGGCCAG)

High quality sequence stop: 704

POLYA=yes.

FEATURES

source

1. 731

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:7474350"

/sex="male"

/tissue_type="embryonic stem"

/cell_type="human embryonic stem cells"

/cell_line="BG01"

/lab_host="DH10B-T1 phage-resistant E. coli"

/clone_lib="NIH_MGC_262"

/note="Vector: pExpress-1; Site_1: NotI; Site_2: EcoRV;

RNA obtained from human embryonic stem cells isolated from

the inner cell mass of blastocyst stage embryos and

differentiated to an early neural progenitor cell type.

Cell line id and NIH Registry designation is BG01.

Positive for Nestin and Musashi expression. Passage number

18. cDNA primed using oligo-dT primer:

5'-pGACTAGTTCTAGATCGGAGCGCCCTT(125)-3' and cloned into

the BcoRV/NotI sites of pExpress-1. This primary library

is non-normalized (normalized primary library is

NIH_MGC_259). It was constructed by Express Genomics

(Frederick, MD). Sequence ends have been trimmed to

exclude vector and regions below phred quality 16. Note:

this is a Mammalian Gene Collection library."

ORIGIN

Query Match 35.1%; Score 707; DB 8; Length 731;
Best Local Similarity 100.0%; Pred. No. 3e-190;

CN389072 CN389072 696 bp mRNA linear EST 16-MAY-2004
LOCUS 17000424036093 GRN_ES Homo sapiens cDNA 5', mRNA sequence.
ACCESSION CN389072 GI:47376667
VERSION CN389072.1
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominoidea; Homo.
REFERENCE 1 (bases 1 to 696)
AUTHORS Brandenberger, R., Wei, H., Zhang, S., Lei, S., Muraige, J., Fisk, G. J.,
Li, Y., Xu, C., Pang, R., Guejler, K., Rao, M. S., Mandalam, R.,
Lebowicki, J. and Stanton, L. W.
TITLE Transcriptome characterization elucidates signaling networks that
control human ES cell growth and differentiation
JOURNAL Nat. Biotechnol. 22 (6), 707-716 (2004)
PUBMED 15146197
COMMENT Contact: Brandenberger R
Regenerative Medicine
Geron Corporation
230 Constitution Drive, Menlo Park, CA 94025, USA
Tel: 650 473 8658
Fax: 650 473 7760
Email: rbrandenberger@geron.com
Insert length: 696 Std Error: 0.00.
Location/Qualifiers
FEATURES
source 1..696
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/cisue_type="embryonic stem cells, cell lines H1, H7, and
H9"
/clone_lib="GRN_ES"
/note="Oligo dt primed, full-length enriched cDNA library
from undifferentiated hES cell lines H1 (p32), H7 (p29),
and H9 (p26) maintained in feeder-free conditions"
ORIGIN
Query Match 33.0%; Score 664; DB 7; Length 696;
Best Local Similarity 99.8%; Pred. No. 5.8e-178;
Matches 675; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
QY 740 AGCCCTACGATGTAAGAAGCTGGGAGGCTTTGATGGGCTCAAGCCTGTTATTC 799
DB 21 AGCCCTACGATGTAAGAAGCTGGGAGGCTTTGATGGGCTCAAGCCTGTTATTC 80
QY 800 ATTAAGAGATTCTACTGTGTGAAAAAACCCTATGATGTAAAGACTGTGAAAAGCCCTTTC 859
DB 81 ATTAAGAGATTCTACTGTGTGAAAAAACCCTATGATGTAAAGACTGTGAAAAGCCCTTTC 140
QY 860 GGGGTGTGATGAGCTCAGCAGACCAAGATTTCCACATGGGGAGAAAGACTTAAGAT 919
DB 141 GGGGTGTGATGAGCTCAGCAGACCAAGATTTCCACATGGGGAGAAAGACTTAAGAT 200
QY 920 GCAAGAAGCTGTGGGAAGACCTTTAGCCGTGTATTAACCTTATTCAGCAAGAAGATTC 979
DB 201 GCAAGAAGCTGTGGGAAGACCTTTAGCCGTGTATTAACCTTATTCAGCAAGAAGATTC 260
QY 980 ATAGTGGGAGAAAGCCTTACGAGTGTAAAGACTGTGGGAAGCCTTTATTTGTGTTCAA 1039
DB 261 ATAGTGGGAGAAAGCCTTACGAGTGTAAAGACTGTGGGAAGCCTTTATTTGTGTTCAA 320
QY 1040 GCTCATTCAGCAATAAAGAAATTCACAGAGTGAAGAAACCTTATGATGTCAAGATGTG 1099
DB 321 GCTCATTCAGCAATAAAGAAATTCACAGAGTGAAGAAACCTTATGATGTCAAGATGTG 380
QY 1100 GGAAGGCTTTATTCGAGCAATTAACCTTA-CTCAGACATCAAGAAATCCACACGGGTGAG 1158
DB 381 GGAAGGCTTTATTCGAGCAATTAACCTTAAGCTCAGATCAAGAAATCCACACGGGTGAG 440
QY 1159 AAGCTTACGATGTAAAGAGTGTGGGAAGCCTTTGCTGGGGTTGAGCCTTCGTTAAG 1218

DB 441 AAGCTTACGATGTAAAGAGTGTGGGAAGCCTTTGCTGGGGTTGAGCCTTCGTTAAG 500
QY 1219 CACGAGAGATATACATACGGGGAGAAAGCCGTATCAAGTGCACAGAAATGTGGAAAGCCCTTC 1278
DB 501 CACGAGAGATATACATACGGGGAGAAAGCCGTATCAAGTGCACAGAAATGTGGAAAGCCCTTC 560
QY 1279 AATTGTGCTATTCACCTCAGTCAAGCAGAGAAATCCAACAGGGGAAACCCCGTATAAA 1338
DB 561 AATTGTGCTATTCACCTCAGTCAAGCAGAGAAATCCAACAGGGGAAACCCCGTATAAA 620
QY 1339 TGTAGAGAGTGTGGGAAGCCTTTCTTTATGATGTGACCTTCGTAAACATGAGGAATT 1398
DB 621 TGTAGAGAGTGTGGGAAGCCTTTCTTTATGATGTGACCTTCGTAAACATGAGGAATT 680
QY 1399 CATACCGGGGTGAAC 1414
DB 681 CATACCGGGGTGAAC 696
RESULT 14
DQ045795 709 bp DNA linear GSS 02-JUN-2005
LOCUS Pan troglodytes ZNF463 gene, VIRUTAL TRANSCRIPT, partial sequence,
DEFINITION genomic survey sequence.
ACCESSION DQ045795
VERSION DQ045795.1 GI:66897010
KEYWORDS GSS.
SOURCE Pan troglodytes (chimpanzee)
ORGANISM Pan troglodytes
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominoidea; Pan.
REFERENCE 1 (bases 1 to 709)
AUTHORS Nielsen, R., Bustamante, C., Clark, A. G., Gianowski, S., Sackton, T. B.,
Hubisz, M. J., Fedel, A. L., Tanenbaum, D. M., Civejlo, D.,
White, T. J., Snijsky, J. J., Adams, M. D. and Cargill, M.
TITLE Direct Submission
JOURNAL Submitted (05-MAY-2005) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
COMMENT This sequence was made by sequencing genomic exons and ordering
them based on alignment. Translation starts at the beginning of
alignment.
FEATURES
source 1..709
/organism="Pan troglodytes"
/mol_type="genomic DNA"
/db_xref="taxon:9598"
gene <1..>709
/gene="ZNF463"
/locus_tag="HC11559"
ORIGIN
Query Match 32.9%; Score 663.2; DB 11; Length 709;
Best Local Similarity 93.8%; Pred. No. 9.8e-178;
Matches 665; Conservative 0; Mismatches 44; Indels 0; Gaps 0;
QY 692 GTTGGGGCAATGACCTTACTCAACATCAAAAATTCATCTGTGGGAGAGCCCTACGAAT 751
DB 1 GTTGGGGCAATGACCTTACTCAACATCAAAAATTCATCTGTGGGAGAGCCCTACGAAT 60
QY 752 GTTAAGACTGTGGGAAGCCTTTTCGATGGGGCTCAAGCCTGTTATTCATTAAGAGATTC 811
DB 61 GTTAAGACTGTGGGAAGCCTTTTCGATGGGGCTCAAGCCTGTTATTCATTAAGAGATTC 120

THIS PAGE BLANK (USPTO)


```

Db      |||||
670  GTTGTGTCGTTTCGCGCAGTAGCCATGACTTTTCTCAGAGGAGGTGGGCGTGTCTCA 729
Qy      |||||
257  ACTGCTCAGAGGGGCTGTGCTGGGAGGTGATGCTGAGAACTACAGTAACCTTGSTCT 316
Db      |||||
730  ACTGCTCAGAGGGGCTGTGCTGGGAGGTGATGCTGAGAACTACAGTAACCTTGSTCT 789
Qy      |||||
317  CACTGGATTGAGAGTCAGATATGAAATAGAGTTTACTTACAGAAAAAAACATTGATG 376
Db      |||||
790  CACTGGATTGAGAGTCAGATATGAAATAGAGTTTACTTACAGAAAAAAACATTGATG 849
Qy      |||||
377  AAATAAGGGCTTCCAAAAGGAATTCAGATAGAAAGTAATCCCTTGCGCGTAATGCA 436
Db      |||||
850  AAATAAGGGCTTCCAAAAGGAATTCAGATAGAAAGTAATCCCTTGCGCGTAATGCA 909
Qy      |||||
437  TAATGGAAGGTACGCTTGAAGACACAGCGCTCCAGAGGAGGTATGTAATCAGATGA 496
Db      |||||
910  TAATGGAAGGTACGCTTGAAGACACAGCGCTCCAGAGGAGGTATGTAATCAGATGA 969
Qy      |||||
497  TCATCAATTATGTCAAAAAGCGCTGTACTAGAGAAAGCAACCCCTCCAGAACATCAGA 556
Db      |||||
970  TCATCAATTATGTCAAAAAGCGCTGTACTAGAGAAAGCAACCCCTCCAGAACATCAGA 1029
Qy      |||||
557  GACATCATTAAGAGAAATTCCTTTGAAATGTAGAGACTGTGGAAAGCCCTTATGTCGT 616
Db      |||||
1030  GACATCATTAAGAGAAATTCCTTTGAAATGTAGAGACTGTGGAAAGCCCTTATGTCGT 1089
Qy      |||||
617  ATCACTTGTAGCAACATCAGAAAATTCATATCTGAGTGAAGAAACCTTTATGAAATGA 676
Db      |||||
1090  ATCACTTGTAGCAACATCAGAAAATTCATATCTGAGTGAAGAAACCTTTATGAAATGA 1149
Qy      |||||
677  GTAAGAGGCTTCCGTTGGGGCAATCAGCTTACTCAACATCAAAAATTCATATCTGGGG 736
Db      |||||
1150  GTAAGAGGCTTCCGTTGGGGCAATCAGCTTACTCAACATCAAAAATTCATATCTGGGG 1209
Qy      |||||
737  AGAAGCCCTTACGAATGTAAAGACTGTGGGAAGGCTTTTGATGGGCTCAAGCTTGTTA 796
Db      |||||
1210  AGAAGCCCTTACGAATGTAAAGACTGTGGGAAGGCTTTTGATGGGCTCAAGCTTGTTA 1269
Qy      |||||
797  TTCATTAAGAGATTCATCTGCGTGAAGAAACCCATGAATGTAAAGCTGTGGAAAGGCT 856
Db      |||||
1270  TTCATTAAGAGATTCATCTGCGTGAAGAAACCCATGAATGTAAAGCTGTGGAAAGGCT 1329
Qy      |||||
857  TTCGCGGTGTGATGAGTCACTCAGCACCAGAGATTTCCACTGTGGGAGAAAGACTAG 916
Db      |||||
1330  TTCGCGGTGTGATGAGTCACTCAGCACCAGAGATTTCCACTGTGGGAGAAAGACTAG 1389
Qy      |||||
917  AATGCAAGAAGCTGTGGGAAGACTTTAGCCGTGTATTAACCTTATTCAGCACAAGAA 976
Db      |||||
1390  AATGCAAGAAGCTGTGGGAAGACTTTAGCCGTGTATTAACCTTATTCAGCACAAGAA 1449
Qy      |||||
977  TTCATAGTGGGAGAAAGCCTTACAGATGTAAAGACTGTGGGAAGGCTTTATTTGTTG 1036
Db      |||||
1450  TTCATAGTGGGAGAAAGCCTTACAGATGTAAAGACTGTGGGAAGGCTTTATTTGTTG 1509
Qy      |||||
1037  CAAGCCTCATTCAGATTAAGAAATTCACACAGAGTGAAGAAACCTTATGATGTCAAGAA 1096
Db      |||||
1510  CAAGCCTCATTCAGATTAAGAAATTCACACAGAGTGAAGAAACCTTATGATGTCAAGAA 1569
Qy      |||||
1097  GTGGGAAGGCTTTACTCGAGTCAATTAACCTTACTCAGCATCAGAAAGTCAACACGGTG 1156
Db      |||||
1570  GTGGGAAGGCTTTACTCGAGTCAATTAACCTTACTCAGCATCAGAAAGTCAACACGGTG 1629
Qy      |||||
1157  AGAAGCTCAACGAATGTAAAGAGTGTGGGAAGGCTTTGCGTGGGGTTCAGGCTTGTA 1216
Db      |||||
1630  AGAAGCTCAACGAATGTAAAGAGTGTGGGAAGGCTTTGCGTGGGGTTCAGGCTTGTA 1689
Qy      |||||
1217  AGCACAAGAGATACATACGGGCGAGAAAGCGTACAAAGTCAAGATGTGGGAAGGCTT 1276
Db      |||||
1690  AGCACAAGAGATACATACGGGCGAGAAAGCGTACAAAGTCAAGATGTGGGAAGGCTT 1749
Qy      |||||
1277  TCAATTGTGCTATCACTCACTCAGACAGAGAAATCCACAGGCGAAACCCCGTATTA 1336

```

```

Db      |||||
1750  TCAATTGTGCTATCACTCACTCAGACAGAGAAATCCACAGAGCGAAACCCGTATTA 1809
Qy      |||||
1337  AATGTAAAGAGTGTGGGAAGGCTTTCAATTTATGATTCAGAGCCTCGTGAACATGAGAGA 1396
Db      |||||
1810  AATGTAAAGAGTGTGGGAAGGCTTTCAATTTATGATTCAGAGCCTCGTGAACATGAGAGA 1869
Qy      |||||
1397  TTCATACCGGGGTGAACCCCTATGSGGTGTACAGAAATGTGGGAAGGCTTTAGTCACGGCC 1456
Db      |||||
1870  TTCATACCGGGGTGAACCCCTATGSGGTGTACAGAAATGTGGGAAGGCTTTAGTCACGGCC 1929
Qy      |||||
1457  ATCAGCTTACACAATCAGAAAAGGCAACGTGGGGCCGAAATCCATCAGATTAAGAGAGT 1516
Db      |||||
1930  ATCAGCTTACACAATCAGAAAAGGCAACGTGGGGCCGAAATCCATCAGATTAAGAGAGT 1989
Qy      |||||
1517  GCGGGAAGGAGTATACCACTTAACCAATCTCCGAGAAATCAGAGATCCACAACAGTT 1576
Db      |||||
1990  GCGGGAAGGAGTATACCACTTAACCAATCTCCGAGAAATCAGAGATCCACAACAGTT 2049
Qy      |||||
1577  GAAAGCCTTTTGAACGAGTAGCCCGCTCGTATCTATGTGTTTGCCTTCCACAGTTTGT 1636
Db      |||||
2050  GAAAGCCTTTTGAACGAGTAGCCCGCTCGTATCTATGTGTTTGCCTTCCACAGTTTGT 2109
Qy      |||||
1637  TACCTGCAGTCAACGTCAGTTCAAAAATTTAAATGGAATAATCCAGAAATTAAGAAATT 1696
Db      |||||
2110  TACCTGCAGTCAACGTCAGTTCAAAAATTTAAATGGAATAATCCAGAAATTAAGAAATT 2169
Qy      |||||
1697  TAAGTCTCAA 1707
Db      |||||
2170  TAAGTCTCAA 2180

```

RESULT 2

```

US-11-112-908-55
; Sequence 55, Application US/11112908
; Publication No. US20050260659A1
; GENERAL INFORMATION:
; APPLICANT: Harris, Cole
; TITLE OF INVENTION: Breast Cancer Biomarkers
; FILE REFERENCE: 04-164-US
; CURRENT APPLICATION NUMBER: US/11/112, 908
; PRIOR FILING DATE: 2005-04-22
; PRIOR APPLICATION NUMBER: US 60/564, 758
; PRIOR FILING DATE: 2004-04-23
; PRIOR APPLICATION NUMBER: US 60/575, 978
; PRIOR FILING DATE: 2004-06-01
; PRIOR APPLICATION NUMBER: US 60/631, 702
; PRIOR FILING DATE: 2004-11-30
; PRIOR APPLICATION NUMBER: US 60/633, 826
; PRIOR FILING DATE: 2004-12-07
; NUMBER OF SEQ ID NOS: 511
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 55
; LENGTH: 193789
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-112-908-55

```

Query Match 23.9%; Score 482.4; DB 8; Length 193789;

Best Local Similarity 65.9%; Pred. No. 8,2e-142;

Matches 699; Conservative 0; Mismatches 361; Indels 0; Gaps 0;

```

Qy      |||||
528  AGAAGCACCCCTCCAGAAACATCAGAGATCACTAAGAGAAATTCCTTGAATGTAA 587
Db      |||||
13944  AATGTGAAGGCTTACCCGAGATCAGAGTACTCACTAGTGAAGAAACCTTTGGGTGTAA 14003
Qy      |||||
588  GGACTGTGGGAAGGCTTTAGTCTGGCTATCAACTTACTCAACATCAGAAAAATCCATAC 647
Db      |||||
14004  GGAAGTGTGGGAAGGCTTTAGTCTGGCTATCAACTTACTCAACATCAGAAAAATCCATAC 14063
Qy      |||||
648  TGGTGAAGAACTTATGAATGTAAAGATGTAAAGAGGCTTCCGTTGGGGCAATACGCT 707
Db      |||||
14064  CAGTGAAGAAACCTTACGAATGTGAAGAAAGATGTGGGAAGGCTTTAGCATAGCTCAACCCCT 14123

```


LENGTH: 4305
 TYPE: DNA
 ORGANISM: Bovine 19866880546157
 US-10-750-623-44166

Query Match 22.1%; Score 444.6; DB 7; Length 4305;
 Best Local Similarity 64.3%; Pred. No. 7.8e-131;
 Matches 699; Conservative 0; Mismatches 384; Indels 4; Gaps 2;

```

QY 487 AATCAGATGATCATCATTAATATGTCAAAAGGCTGTCTACTAGAGAGGCAACCCCTCTAGA 546
DB 3523 AACCATATGAGTGAAGACTGTGGAGAGGCTTCAACCAAGAGTCAACACTTGTCTCAGC 3464
QY 547 AACATCAAGACATCATTAAGAGAAATTCCTTGAATGAAGACTGTGGAGAGGCTTT 606
DB 3463 AC-CACAGAAATACACCCGAGAAAAATTTGAATGAAGAAATGAAGAAAGCTTTC 3405
QY 607 AGTCGTGGCTATCAACTTGTCAACATCAAGAAAAATCCATCTGTGAGAAACCTTATGAA 666
DB 3404 AGCCAAAAATGTACATCTTATTCAACATCAAGAAATTCATCTGAGAAAAACCATATTA 3345
QY 667 TGTAAAGAAATGAAGAGGCTTTCGTTGGGCAATCACTTACTCAACATCAAAAAAT 726
DB 3344 TGTAAAGAGTGAAGAAAGCTTTCAGCCAGCTGACACCTTGTCTCAGCATCAGAAAT 3285
QY 727 CATCTGGGGAGAAAGCCCTACGAATTAAGACTGTGGAGAGGCTTTGATGGGCTCA 786
DB 3284 CATACGGGGAGAAAGCTTATTAATTAAGAAATGTGGAGAGGCTTCAAGTATGCTCG 3225
QY 787 AGCCTGTTATTCATTAAGAGATTCATCTGTGAAAAACCTTGAATGAAGACTGT 846
DB 3224 TCCTTTGGCCGACATAGAGGTTCACCTGGCAAAAGACCTTGAAGATTTAGATGT 3165
QY 847 GGAAGAGCTTTCGGGCTGTGATGAGCTCACTCAGCAACCAAGA---TTCACACTGGG 903
DB 3164 GGGAAAGCTTTCAGGCAAAACATATGCTTATCCGTCACTGAGATTAATATCACTGGG 3105
QY 904 GAGAAAGACTAGAAATGCAAAAGCTGTGGAGAGCTTTAGCCGTGTGATTAACCTTAT 963
DB 3104 GAGAAAGCTTTCAGCTGATCACTGTGGCAAGGCTTTCAGGCATCACTAGACTTAT 3045
QY 964 CAGCACAAGAGATTCATGTGGGAGAAAGCTTTCAGATGAAGACTGTGGAGAGCT 1023
DB 3044 CAGCACAAGAGATTCATGTGGAGAAACCTTCAAAATGATATGTGTGGAGAAAC 2985
QY 1024 TTTATTTGTGTTCAAGCTTATTCAGCATTAAGAAATTCACAGGTGAGAAACCTTAT 1083
DB 2984 TTCAGTACGGCTCATCCCTGACCGTCAATCAGAAATTCACAGAGAGAAACATAT 2925
QY 1084 GAATGTCAAGATGTGGAGAGGCTTTACTCGAGTCAATTAACCTTCTCAGCATCGAAG 1143
DB 2924 GAATGTCAATCTGTGGAGAAAGCTTTCAGGCATATGCTCTCACCAGCAACAA 2865
QY 1144 ATCCACACCGGTGAGAAAGCTTCAGATGTGAAGAGTGGAGAGGCTTTCGCTGGGT 1203
DB 2864 GTGCATCTGTGAGAGAGGCTTTCAGATGTGGAGAAATGTGGAGAAACCTTTCAGGCA 2805
QY 1204 TCGAGCCTGTTAAGCAAGAGAGATTAACGCGGAGAGAGGCTTTCAGATGCAAGAA 1263
DB 2804 ATACACCTTGTGATGATCTGAGGATTCATATCGAGAGAAACCTTATGAATGAAGAA 2745
QY 1264 TGTGGAGAGGCTTCAATGTGCTATCACTCACTCAGACGAGAGAAATCCACAGAGC 1323
DB 2744 TGTGGAGAGGCTTTCAGATGATCACTGAGCTGATCATCATGAGAAATTCATCTGA 2685
QY 1324 GAAACCCCTATTAATGAAGAGTGTGGAGAGCTTTTCAATTAATGATCGAGCTGTG 1383
DB 2684 GAGAAACCTTATCAATGTGAAGAGTGTGGAGAGGCTTTCAGACAGAGGCACTCTGCA 2625
QY 1384 AAACATGAGAAATTCATACCGGGGTGAACCTTATGGGTGATCAGAAATGTGGAGAGC 1443
DB 2624 CAGCACAATTAATCATCTGAGAGAGAAACCTTATGATGTATGAATGTGTAAGCC 2565

```

```

QY 1444 TTTAGCAGGCGCATCAGCTTACACAACTCAGAAAAACGACAGTGGGCGAAATCTTAC 1503
DB 2564 TTCAAGCCAGGCTTACCCGCTTATTCAACTCAGAGATTTCAACAGAGAGAAAGCTTAT 2505
QY 1504 GAATGTAAAGAGTGTGGAGAGGCTATGTAACAATAACATCTCCGAGAACATCAGAG 1563
DB 2504 GAATGTCAATGATGTGGAGAGGCTTTAGTGAAGCTCATCCGTCGATCAGATCGAGA 2445
QY 1564 ATCCACA 1570
DB 2444 CTTCACA 2438

```

RESULT 6 US-10-750-185-49223

```

; Sequence 49223, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: NMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INHERRING BOVINE TRAITS
; FILE REFERENCE: MM11100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; PRIOR FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 49223
; LENGTH: 1730
; TYPE: DNA
; ORGANISM: Bovine 19866880440784
; US-10-750-185-49223

```

Query Match 21.8%; Score 439.6; DB 7; Length 1730;
 Best Local Similarity 63.2%; Pred. No. 1.7e-129;
 Matches 707; Conservative 0; Mismatches 409; Indels 2; Gaps 2;

```

QY 528 AGAAGGCAACCTCTCTAGAAACATCAGAGA-CATCATTAAGAGAAATTCCTTGAATGA 586
DB 130 AAGATTCATCCCTTGTCTCAAAACATGAGAAATCATTCGAGAGAGCCCTTGAATGA 189
QY 587 AGAAGTGTGGAGAGGCTTGTAGTGTGCTATCACTTAAGTCAACATCAGAAATCATATA 646
DB 190 ATCAAGTGTGAAATTCCTTCAAGTGTGAGCTTCATCTTGTGCAATCAGAGAACTCAT 249
QY 647 CTGTGTGAGAAACCTTATGAATGTAAAGAAATGTGAAGAGGCTTTCGTTGGGCAATCAGC 706
DB 250 CAGGGGAGAAACCTTATGAATGTAAATGTGAAGTGGAGAAATCATTCAGCCGACCTTCACC 309
QY 707 TTACTCAACTCAAAAAATTTCTA-CTGGGAGAGAGCCTTAAGAAATGAAGACTGTGG 765
DB 310 TTGTTCCTCATGAAGAACTCATCTGAGAGAAACCTTACAGATGTATCAGTGTGT 369
QY 766 AAGGCTTTTGAATGGGCTCAAGCCTGTTATTCATTAAGAGATTCATCTGATGAAGAA 825
DB 370 AATCTTTTGAAGAGATGTGTAAGAGGCTTTCGCGGTGTGATGAGCTCACTCAGAC 885
QY 826 CCTATGAATGTAAAGATGTGAAAGGCTTTCGCGGTGTGATGAGCTCACTCAGAC 885
DB 430 CCTTACAGATGCAATCAGATGTGAAAGGCTTTCAGACAGAGCTCAACAACTTATCGCACAT 489
QY 886 CAGAGATTCACACTGTGGAGAGAAAGCTACGAATGGAAGACTGTGGAGAGAGCCTTAC 945
DB 490 CAAAGAACTCATCTGAGAGAGAGCCCTTACGAATGTAAATGTGGAGAAATCATTTATC 549
QY 946 CGTGTGTATTAATCTTATTCAGCAAGAGAAATTCATAGTGGGAGAGGCTTACGAGTGT 1005

```

```
Db      550 CAGAGCTATAAATTATGCCCCATCAAGAATTCATACGTGAGAAAAACCTATGAAATGC 609
QY      1006 AAAGACTGTGGAGGCTTTTATTTGNGTTCAAGCCTCATTCAGCATAAAGAAATTCAC 1065
Db      610 AATCAATGTGGAAAGCTTTTAACTCAAAAGTTACAAACTTGTCCACCCAGAAATCTCAC 669
QY      1066 ACAGGTGAGAAACCTATGAAATGCAAGATGTGGAAAGCCTTTTACCTGAGTCAATTAC 1125
Db      670 ACAGGAGAAAAACCTTTGAAATGTAATCAAGTGTGAAAAATCTTCAGCTGAGACTCTCAA 729
QY      1126 CTTACTCAGCATAGAAATCCACACCGGTGAGAAACCTTCAGAAATGTAAGAGTGTGG 1185
Db      730 CTGTGTGCAACCAAGAACTCACCTGTGAGAGAAACCTTACGAAATGTAATGATGTGG 789
QY      1186 AAGGCTTTTCGTGGGGTTCGAGCCTGTATTAACAAGAGATATACATACGCGCGAGAAAG 1245
Db      790 AATCTTTTCAACCGCAGTTCTCACCTTTATGATCAAGAACTCACACTGAGAGAAAA 849
QY      1246 CCGTACAGTGACAGAAATGTGGAAAGCCTTCAATTTGTGCTATACCTCACTCAGCAC 1305
Db      850 CCTATGAATGTAAACAGTGTGGAAAGTCTTCAGCAGAGTTATGTCTTGTGTACAT 909
QY      1306 GAGAGAAATCCACACAGCGGAAACCCGTAATAATGTAAGAGTGTGGAAGCTTTCATT 1365
Db      910 CAGAGAACTCACCTGCGGAAAGCCCTATGATGACGTGATGTGAAAAATCTTCAGG 969
QY      1366 TATGATTCGAGCCTCTGAAACATGAGAAATTCATACCGGGGTGAAACCTATGGGTGT 1425
Db      970 CAGAGTTCAATGCTTACTCAACATCAGAGAACTCATCTGAGAGAAACCTTATGAAATGT 1029
QY      1426 ACAGAAATGTGGAGAGCTTTAGTACAGCGGCATCAGTTACACACATCAGAAACGCAAC 1485
Db      1030 AATCAATGTGGAAAAACATTCACGTTGAGTGTGCACTTATGTATCATCAAGAACTCAT 1089
QY      1486 AGTGGGCGCAAAATCCTACGAATTTAAGAGTGTGGGGAAGGCACTTACCACTAAACCAT 1545
Db      1090 ACTGAGAGAAACCTTTAGTGTAAATCAATGTGGAAAGCTTTCATTAATGTTCTTAA 1149
QY      1546 CTCGGAACATCAGAGATCCACAACAAGTTGAAGCCTTTGAAACGAGTACCGGCT 1605
Db      1150 CTATATAGACATAGGCAACTCATCAGAGAGAAACCTTATGATGTAACTGACCTGGA 1209
QY      1606 CGTATCTAATGTTTCGCTTTCACAGTTGTACTGTC 1643
Db      1210 AATCTTTCAGAGAGATATCCTTCCTCCTGCTCC 1247
```

RESULT 7

```
US-10-750-623-49223
; Sequence 49223, Application US/10750623
; Publication No. US20050287531A1
; GENERAL INFORMATION:
; APPLICANT: MMT GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFIELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: PANTIN, Dennis
; TITLE OF INVENTION: METHODS AND SYSTEMS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MW11100-1
; CURRENT APPLICATION NUMBER: US/10/750,623
; PRIORITY FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: Patent version 3.1
; SEQ ID NO 49223
; LENGTH: 1730
; TYPE: DNA
; ORGANISM: Bovine 19866880440784
US-10-750-623-49223
```

```
Query Match      21.8%; Score 439.6; DB 7; Length 1730;
Best Local Similarity 63.2%; Pred. No. 1.7e-129;
Matches 707; Conservative 0; Mismatches 409; Indels 2; Gaps 2;

QY      528 AGAAGGCAACCCCTCTAGAAACATCAGAGA-CATCATAGAGAAATTCCTTGAATGA 586
Db      130 AAGATTCAATCCCTTCTCAAAACATGAGAAATCATTCGAAAGAAAGCCCTTTGAAATGA 189
QY      587 AGGACTGTGGAAAGCCTTTAGTCGTGCTATCAACTTATAGTCAACATCAGAAATTCATA 646
Db      190 ATCAAGTGTGAAATTCCTTCACTTGAAGCTCTCACTTGTTCACATCAGAGAACTCAT 249
QY      647 CTGTGTGAAACCTTATGAAATGTAAGAAATGTAAGAGGCTTCCGTTGGGCAATCAGC 706
Db      250 CAGGGGAGAAACCTTATGAAATGTAATGTAATGTGGGAAATCATTCAGCCCACTCACCC 309
QY      707 TTAATCAATATAAAATATTCATA-CTGGGGGAGAGCCCTAGAAATGTAAGAACTGTGG 765
Db      310 TTGTTTCCCATCAAGAACTCATACCTGAGAGAAACCTTACAGATGTAATCATGATGTGT 369
QY      766 AAGGCTTTTCAGATGGGGCTCAAGCCTGTTATTCATTAAGAGATTCATACGTGTGAAAA 825
Db      370 AATCTTTTACGCAAGCTATATGCTTGTGTGATCAGAGAACTCATCTGAGAGAAAG 429
QY      825 CCTATGAATGTAAAGCTGTGAAAGGCTTTCCGCGTGTGATGACTCATCAGCAC 885
Db      430 CCTTACGAGTGAATCAGTGTGAAAGTCAATTCAGACAGCTACAACTTATTCGCAAT 489
QY      886 CAGAGATTCACACTGTGGGAGAAAGCTACGAATGAAAGACTGTGGGAAAGCCTTTAGC 945
Db      490 CAAGAACTCATCTGAGAGAAAGCCCTACGAATGTAATCAATGTGGGAAATCATTTATC 549
QY      946 CCGTGTATATAACTTATTCAGACAAGAAATTCATAGTGGGAGAAAGCTTACGAGTGT 1005
Db      550 CAGAGTATAAATCTTATTTGCCATCAAGAAATTCATCTGAGAGAAAGCCCTATGAAATGC 609
QY      1006 AAGACTGTGGAGGCTTTTATTTGTGTCAAGCTTCATTCAGATATAAAGATTCAC 1065
Db      610 AATCAATGTGGAAAGCTTTTATGTAAGAAATTCAAATCTTTGTCACCAACAGAACTCAC 669
QY      1066 ACAGGTGAGAAACCCCTAGAAATGTCAAGAAATGTGGGAAAGCCTTATCTGAGTCAATTC 1125
Db      670 ACAGGAGAAAAACCTTGAATGTAATCAGTGTGAAAAATCCTTCAGCTGAGACTCTCAA 729
QY      1126 CTTACTCAGCATCAGAAATCCACAACCGGTGAGAAACCTTCAGAAATGTAAGAGTGTGG 1185
Db      730 CTGTGTGCAACCAAGAACTCACCTGTGAGAGAAACCTTATGAAATGTAATGATGTGG 789
QY      1186 AAGGCTTTTCGTGGGGTTCGAGCCTGTATTAACAAGAGATATACATACGCGCGAGAAAG 1245
Db      790 AATCTTTTCAACCGCAGTTCTCACCTTTGTATGCAATCAGAGAACTCACCTGAGAAAAA 849
QY      1246 CCGTACAGTGACAGAAATGTGGAAAGCCTTCAATTTGTGCTATACCTCACTCAGCAC 1305
Db      850 CCTATGAATGTAAACAGTGTGGAAAGTCTTCAGCAGAGTTATGTCTTGTGTACAT 909
QY      1306 GAGAGAAATCCACACAGCGGAAACCCGTAATAATGTAAGAGTGTGGAAGCCTTTCATT 1365
Db      910 CAGAGAACTCACCTGCGGAAAGCCCTATGATGACGTGATGTGAAAAATCTTCAGG 969
QY      1366 TATGATTCGAGCCTCTGAAACATGAGAAATTCATACCGGGGTGAAACCTTATGGGTGT 1425
Db      1426 ACAGAAATGTGGAGAGCTTTTATGATCAGCGGCATCAGCTTACACAAATAGAAAGGCAAC 1485
QY      1486 AGTGGGCGAAATCCTACGAATGTAAGAGTGTGGGGAAGGCACTTACCACTTAAACCAT 1545
Db      1090 ACTGAGAGAAACCTTTACGTGTATTCAAATGTGGGAAAGCTTTCATTAATGTTCTTAA 1149
QY      1546 CTCGGAACATCAGAGATCCACAACAAGTTGAAGGCTTTTGAAGGAGTACCGGCT 1605
```

```

Db      1150 CTTATTAGCATCAGGCACTCATACAGAGAGAAACCTTAGATGATTAAGTCTGGA 1209
Qy      1606 CGTATCTATGTTTGGCTTTTCACAGTTTGTACCTGC 1643
Db      1210 AATCTTTGACAGAGATATCCCTTCCTCTGCTGCC 1247

RESULT 8
US-10-775-169-119
; Sequence 119, Application US/10775169
; Publication No. US20050287532A9
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Burczynski, Michael
; APPLICANT: Twine, Natalie
; APPLICANT: Dorneier, Andrew
; APPLICANT: Trepicchio, William
; TITLE OF INVENTION: Method for Monitoring Drug Activities In Vivo
; FILE REFERENCE: AM101080 (031896-013000)
; CURRENT APPLICATION NUMBER: US/10/775,169
; NUMBER OF SEQ ID NOS: 5278
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO: 119
; LENGTH: 5871
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-775-169-119

Query Match      20.9%; Score 421.4; DB 7; Length 5871;
Best Local Similarity 62.4%; Pred. No. 2.4e-123;
Matches 676; Conservative 0; Mismatches 406; Indels 1; Gaps 1;

Qy      506 ATGTCAAAAGGCTGCTACTAGAGAGAGCCCTTCTAGAACATCAGACATCATATA 565
Db      508 AAGTCAGAGAGAAACCATATTTCTGAACCTTCTTCTGCTTCAAGAGAGCTCATTA 567

Qy      566 -AGAGAAATTCCTTGAATGTAGAGACTGTGGAGAGCCCTTATGTCGTATCAACTT 624
Db      568 GAGGGGAGAACCAATCCTTGCGATATCTGTGGCAAAACCTTACGTTTATTCGAGCTTA 627

Qy      625 AGTCAACATCAGAAAATTCATCTGTGTAGAAACCTTATGAATGTAAAGATTAAGAG 684
Db      628 GTTAGGCAATCGAATTTCTGATCTGTGGAGAAACCTTACACGTGCGATCAGTGTGGAAA 687

Qy      685 GCCTTCGTTGGGCAATCAGCTTACTCAACATCAAAAATTCATCTAGTGGGAGAGAGCC 744
Db      688 GCGTTTGCCCAAGCTCAGACCTTATGAGCATCAGAGAAATTCACACTGGAAGAGACTC 747

Qy      745 TAGCAATGTAAAGACTGTGGAGAGCTTTTTCATGAGGGGCTCAAGCCTGTTATTCATAAG 804
Db      748 TAGCTCTGAATGTGTGTGGAGAAAGCTTCATTTCACTATTACAGTCTCATTGAGCATCAG 807

Qy      805 AGGATTCATACGTGTGCAAAAACCTTATGAATGTAAAGATGTGAGAAAGCCCTTGGGCT 864
Db      808 CCGCTTCATTCAGAGAGAAAGCCCTTCAATGTGCGCAATGTGGAGAGGCGTTTGTGCAC 867

Qy      865 GGTGATGAGCTCAGTCAGACAGAGATTCCAACATGTGGAGAGAAAGATCAAGAAATGCACA 924
Db      868 AGTTGAGACTGATTAAGGACACAGAGAGTTTCAACAGAGAGAGACTTGTGAATGCACA 927

Qy      925 GACTGTGGAGAGACCTTTAGCCGTGTATTAACCTTATTCAGACAGAGAAATTCATAGT 984
Db      928 GAGTGTGGAGAGAGGCTTCAGTCAGAGCTCTTATCTTATTCGCAATCAGAGAGATTTCACAG 987

Qy      985 GGGGAGAGACCTTTAGAGGTAAAGCTGTGGAGAGGCTTTTATTTGTGTTCAGAGCCCTC 1044
Db      988 GGAAGAGAGCCCTTATGAGTGCAATGATGTGGAGAAATCTTCAITTAAGAGAGCTCAGAGCTTC 1047

Qy      1045 ATTCAAGCATTAAGATTCACACAGGTGAGAAACCTTATGAATGTCAAGATGTGGAGAG 1104
Db      1048 ATTGCGCATTTATCAATCAACAGAGAGTGAAGAGTATGAAATGCAAGAAATGTGGAGAG 1107

```

```

Qy      1105 GCCTTACTCGAGTCATTTACCTTATCTGAGATCAGAAATGCCAGCCGTGAGAAAGCCT 1164
Db      1108 GCATTCGCTATCTGCTGACAGACCTTATTAAGAACACAGAAATTTACACACGAGAGAGAGCC 1167

Qy      1165 CACGAATGTAAAGAGTGTGGAGAGGCTTTGCTGGGGTTCAGAGCTGTTAAGACAGAG 1224
Db      1168 TTGAAATGCATGATGTGTGGAGAGGCTTTATTCGAGAGTTCAAAGCTCATTCAGATCAG 1227

Qy      1225 AGGATATCATACGAGGAGAGAGCCGTACAGATGACAGAAATGTGGAGAGGCTTCATATGT 1284
Db      1228 AGGATTCATACCTGGAGAGAGGCTTACGATATGATGATGTGGAGAGGCTTCAGAGCAG 1287

Qy      1285 GCGTATCACTTCACTCAGACAGAGAGATTCACACAGGCGAAACCCCGTATTAATAGAG 1344
Db      1288 ACGTAAACCTTCAACCCACATCAGAGAAATTCACACTGAGAGAAACTGTATAAATGTAAAC 1347

Qy      1345 GAGTGTGGAGAGGCTTTATTTATGATCGAGCTCTGTAAACATGAGAGAAATTCATACC 1404
Db      1348 GAGTGTGGAGAGGCTTTCTTTCTGTGATTCATACCTTATTCAGACACAGAAATTCACACT 1407

Qy      1405 GGGGTGAAACCTTATGGGTGTACAGAAATGTGGAGAGCTTTATGTCAGGCGCATCAGCTT 1464
Db      1408 GGAGAGAGAGTGTATGATGTAAAGAAATGTGGAGAGGCTTCTCCAGAAAGCCCATCTTC 1467

Qy      1465 ACACAACTCAGAAAACGACAGTGGGCGAAATCTTACGAATGTAAAGAGTGTGGAGAG 1524
Db      1468 ACTGAGACACAGAAATCATCTGTGGAGAGGCGCTTCTGAATGTAAAGATGTGGAGAG 1527

Qy      1525 GCATGTAAACCACTTAACCATCTCCAGAACATCAGAGAGATCCACAGTTGAAGAGCC 1584
Db      1528 GCTTCATATCAGAGCTCCAGAGCTGCTTCTGACACAGATTTATTCACATCGAGAGAAAGCC 1587

Qy      1585 TTT 1587
Db      1588 TAT 1590

```

```

RESULT 9
US-10-750-185-56682/c
; Sequence 56682, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMT GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: PANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; PRIOR FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO: 56682
; LENGTH: 934
; TYPE: DNA
; ORGANISM: Bovine 19866880596109
US-10-750-185-56682

Query Match      20.5%; Score 413.8; DB 7; Length 934;
Best Local Similarity 68.0%; Pred. No. 1.9e-121;
Matches 591; Conservative 0; Mismatches 277; Indels 1; Gaps 1;

Qy      550 CATCAGACATCATTAAGAGAAATTCCTTTGAATTAAGAGCTGTGGAGAGGCTTTTATG 609
Db      875 CAGAAAATTCATCATGATGTGTGAACCTTATATCTGTAAAGATGTGAAAGGCGCTTTTGGT 816

Qy      610 CGTGCTATCAACTTATGTCACATCAGAAATTCATATCTGTGTGAGAAACCTTATGATGT 669

```



```
RESULT 11
US-11-000-688-1136
; Sequence 1136, Application US/11000688
; Publication No. US20050287544A1
; GENERAL INFORMATION:
; APPLICANT: BERTUCCI, Francois
; APPLICANT: HOUIGATTE, Remi
; APPLICANT: BIRNBAUM, Daniel
; TITLE OF INVENTION: GENE EXPRESSION PROFILING OF COLON CANCER WITH DNA ARRAYS
; FILE REFERENCE: 1423-R-03
; CURRENT APPLICATION NUMBER: US/11/000,688
; PRIOR FILING DATE: 2004-12-01
; PRIOR APPLICATION NUMBER: US 60/525,987
; PRIOR FILING DATE: 2003-12-01
; NUMBER OF SEQ ID NOS: 1596
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1136
; LENGTH: 2435
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial sequences:primer
; NAME/KEY: misc feature
; LOCATION: (1)..(2435)
; OTHER INFORMATION: zinc finger protein 354a(ZNF354A) gene.
US-11-000-688-1136

Query Match      20.2%; Score 407.2; DB 8; Length 2435;
Best Local Similarity 61.7%; Pred. No. 4.6e-119;
Matches 649; Conservative 0; Mismatches 403; Indels 0; Gaps 0;

536 CCCCTCTAGAACACATCAGACATCAAGAGATTCCTTGAATGTAAAGACTGTG 595
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      CCTTCGTAAACATGAGAAAAACATAGAGAGAACTATTAAAGTAAAGATGTT 894
QY      GGAAGGCTTTAGTGTGGCTATCAACTTATGCAACATCAGAAAAATCCATCTGTTGAGA 655
Db      CAAAAGCTTTAGCCAAAGTTCAAGCTCTTATTCACATCAATTAACATCTGAGAGA 954
QY      AACCTTATGAATGTAAAGATGTAAAGAGGCGCTTCGTTGGGGCAATCAGCTTACTCAAC 715
Db      AACCTTATGAATGTAAAGATGTAAAGAGGCGCTTCGTTGGGGCAATCAGCTTACTCAAC 1014
QY      ATCAAAAAATTCATACCTGGGAGAAAGCCCTTACGAATGTAAAGACTGTGGAAAGCCTTTTC 775
Db      ATCTAAGAACCATATCTGTGGAGAAATCTCTACAGATGTAAAGATGTGGTAAATCCTTCA 1074
QY      GATGGGGCTCAAGCCTCGTATTTCTTAAGAGATTCATCTGCTGTAAGAAACCTTATGAT 835
Db      GCGGAAGGTCAAGGCTTTTATATACATCAAAAAATTCATGCTGAAGAAAAACCTTGTAAGT 1134
QY      GTAAGACTGTGGAAAGGCGCTTTCGCGGTGATGAGTCACTGACAGCAGAGATTCC 895
Db      ATATATCCGGGTAGGAAGGATTTAGTTGACAGCAATCCCTTTCTGAGATGTAAAGATTC 1194
QY      AACTGTGGGAGAGAACTACGAATGCAAGACTGTGGAGAGACTTTAGCCCTGTGTATA 955
Db      ATCTAGAAAGAAAGCTTACTTAATGTAAATGTGGAGAACCTTTAAGTACTACTCAT 1254
QY      AACTTATTCAGCAAGAGAAATTCATAGTGGGAGAGGCTTTACGAGTGTAAAGACTGTG 1015
Db      CCTTCGTATATCATGAGAAATTCACACTGAGAGAAAGCCTTTTAAATGTATGATGTG 1314
QY      GGAAGGCTTTATTTTGTGTTCAAGCTCATTCAGCATTAAGAAATTCACAGGTGAGA 1075
Db      GGAGAGCCTTCAGCAGAGATGCTCTCTTATTCACATGAAGAAATTCACACGAGAAA 1374
QY      AACCTTATGATGTCAAGAAATGTGGAAAGGCTTTTACTGAGTCAATTAACCTTACTCAGC 1135
Db      AGCCCTATAGATGATGATGATGTGGAAAGGCTTTTACTTCTATTCACGACTTAATAGAC 1434
```

```
QY      ATCAGAGATCCACACCGGTGAGAAAGCCTCAAGATGTAAAGATGTGGAAAGCCTTTC 1195
Db      ACCGAATCATTCATCTGAGAGAAAGTATTAATTTGATGATGTGTAAAGCCTTAA 1494
QY      GCTGGGGTTTCAGAGCCTCGTTAAGCAGAGAGATACATACGGGCGAGAGCGTACAAAT 1255
Db      GCTCCACTCAACACTTATTTATTCAGAGCGAATTCATATCGAGAAAAACCATGTAAAT 1554
QY      GCACGAATGTGGAAAGGCTTCAATTTGTGCTATCACTCACTAGACAGAGAAATTC 1315
Db      GTAAAGTATGTGGAAAGGCTTTCAGACAGAGATTCAGCTTCAATCAATCAGAGAAATGC 1614
QY      ACACAGGAGAAACCCCGATTAATGTAAAGATGTGGAGAGGCTTCAATTAATGTGCA 1375
Db      ATACTGAGAAAGAGCCTTATTAATGTAAAGATGTGGAGAAACATTCAGGTAACTCAT 1674
QY      GCTCGTGAACATGAGAGAAATTCATACCGGGGTGAAAACCTATGGGTGTACAGAAATGTG 1435
Db      CACTTAGTATCACACGAGAAATTCATCTGAGAGAGAAACATATGATGTAGAGATGTG 1734
QY      GGAAGGCTTTAGTACCGGCAATCAGCTTACACATCAGAAAAACGACAGTGGGCGA 1495
Db      GGATATCTTTTGGCCAAAGTTCAAGCTCTTATTCAGCATGGAAGATTCATACAGAGAAA 1794
QY      AATCCTACGAATGTAAAGAGTGGGAGAGGATGTAAACCATTAACATCTCCGAGAAC 1555
Db      AACCTTTAAATGTATATCATGTGTGAAAACTTTTGAACAAAGCTCATCATGATTTGAC 1854
QY      ATCAGAGATCCACACAGTTGAAGAGCCTTT 1587
Db      ATCAGAGAAATTCATCTGAGAGAAACCTTAT 1886

RESULT 12
US-11-000-688-1179
; Sequence 1179, Application US/11000688
; Publication No. US20050287544A1
; GENERAL INFORMATION:
; APPLICANT: BERTUCCI, Francois
; APPLICANT: HOUIGATTE, Remi
; APPLICANT: BIRNBAUM, Daniel
; TITLE OF INVENTION: GENE EXPRESSION PROFILING OF COLON CANCER WITH DNA ARRAYS
; FILE REFERENCE: 1423-R-03
; CURRENT APPLICATION NUMBER: US/11/000,688
; PRIOR FILING DATE: 2004-12-01
; PRIOR APPLICATION NUMBER: US 60/525,987
; NUMBER OF SEQ ID NOS: 1596
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1179
; LENGTH: 2603
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial sequences:primer
; NAME/KEY: misc feature
; LOCATION: (1)..(2603)
; OTHER INFORMATION: zinc finger protein 3 (zf-51) (ZNF3) gene.
US-11-000-688-1179

Query Match      20.0%; Score 402; DB 8; Length 2603;
Best Local Similarity 60.4%; Pred. No. 2.2e-117;
Matches 737; Conservative 0; Mismatches 470; Indels 13; Gaps 4;

181 AAAACAATGGCCAGAGGTTTGTGACGTTCCGCCGACGTAGACATTTCTTCAGAG 240
Db      AAGCCAGAGTCCAGAGACTGTAACTTTGAGAGATGTAGCTGTATCTTATCCGAAG 459
QY      GAGTGGGCTGTCTGAATCTGTCTGACAGGAGCTGTACTGAGAGTATGTGAGAAC 300
Db      GAGTGAAGCGCTTTGGAACCTGCTCAGAGGAGCCTCTATAGAGATGATGCTGAGAAAT 519
```

```

QY 301 TACAGTAAGTGTCTCACTGAGATTTGAGTCAGCATATGAATAATGAAGTTTACTTACA 360
Db 520 TACGGAAATGTGTCTCACTGAGATCGTAGACGAGACTGAAATATGATCAAGAAATTTCT 579
QY 361 GAAAAAACAATTCAGTAATTAAGGCTTCCAAAAGAAATTCAGATGAAGAAATTAATCC 420
Db 580 GAAAGCA-----CAGATCAATGGGGTCTTACTGGA--AGATTTCAAAAGGAATTTT 631
QY 421 CTTGGCCCGTAATGATATGTGAAGGTAGCTTTGAAGACCAAGGCTCCAGAGGAGAG 480
Db 632 CTCAGGGTCTCAAGTTTAAAGAACCTATGAACGAAATCACTCTGAAAAGCCGCTGG 691
QY 481 TATGTCAATCAATGATCATCAATTATGTCAAAAGGCTGTACTAGAGAGGCAACCTT 540
Db 692 GGAATCTCCCTGGAGAAAGACTGAACAGGAAATATGCAATTTTGGTCAAGTACAGTTG 751
QY 541 CCTAAGAACATCAGAGACATCATTAAGAGAAATTCCTTGAATGTGAAGACTGTGGGAG 600
Db 752 AGAGAGAGCTAACCCCGAGGAGG---GAGAGAGGAGAGAAATATATATATTTGGGAA 807
QY 601 GCCTTTAGTCGTGGCTATCACTTAAGTCAACATCAGAAATCCATATCTGGTGAAGAACT 660
Db 808 AGCTTCACTGTGAATTCACACCTTATCTCAGATCAGAGACTCCCGTGGAGAGACAGACC 867
QY 661 TATGAATGTAAAGATGTAAAGAGCCCTCCGTGGGAGCAATCAGCTTACTCAATCA 720
Db 868 CATTAAGTGTGAATGTAGCAAGAGCTTAAATCGAACTTCAAGACCTTATTCACATCAG 927
QY 721 AAAATTCATACGTGGGAGAGCCCTAGATGTAAAGACTGTGGGAGGCTTTGGATGG 780
Db 928 AGAATCCACATCGGGGAAAGCCCTATGATGTAAAGAGTGGGAGGCTTCACCCAG 987
QY 781 GGTCAAGCCCTGTATTCATTAAGAGATTCATCTGTGTAAGAAACCTATGATGTAA 840
Db 988 AGCTCACACTTATTCAGATCAGAGATTCACACTGGGAGAAACCTTATGATGTAGT 1047
QY 841 GACTGTGAAGAGCCCTTGGGCTGTGTATGAGCTCACTCAGACAGAGATTCACACT 900
Db 1048 GATTTGGGAGAAACCTTCAGCTGTAGCTGTCCCTCATTTCTGATCGAGAGATCCACAG 1107
QY 901 GGGGAGAAAGCTAGAGATGCAAGACTGTGGGAGAAACCTTTAGCCGTGTGTAACTT 960
Db 1108 GGGGAGAAACCTCATTAATGTATGAGTGGGAGAAACCTTCAGCTGGAGCTCCACCTC 1167
QY 961 ATTCAAGCAAGAGATTCATAGTGGGAGAAAGCTTACAGAGTAAAGACTGTGGGAG 1020
Db 1168 ACCCAACCATCAGAGATTCACACTGGTGAAGAAACCTTAGCCCTGCAATGATGTGGAG 1227
QY 1021 GCTTTATTTTGTGTTCAAGCCTCATTCAGCATTAAGAAATTCACAGGTGAAGAACCC 1080
Db 1228 GCCTTCAGAGAGGCTCAACCTTATTCACATCAGAGATTCACACTGGAGAAACCC 1287
QY 1081 TATGAATGTCAAGATGTGGGAGAGGCTTTACTCGAGTCAATTAACCTTACTCGCATCG 1140
Db 1288 TATGAATGTATGAATGTGGGAGAGGCTTCAGCAGAGCTCACACCTTATTCAGCACAG 1347
QY 1141 AAGATCCACAGCGGTGAGAAAGCTCAGGAATGTAAAGAGTGGGAGAGGCTTTCCCTGG 1200
Db 1348 AGAATCACTGAGAGAGAGCCCTTAGATGTATGGAATGTGAGAGAAAGTTTACTTAC 1407
QY 1201 GGTTCAGGCTCGTTAAGCAGAGAGATTCATACGGGAGAGAGCCGTGCAAGTGCACA 1260
Db 1408 AGTTTCAGGCTTATTCAGATCAAGAAATTCACACCGGAGAGAAACCCCTTATGATGTAGT 1467
QY 1261 GAATGTGGAGAGGCTTTCATTTGTGCTATCACTCAGTACAGACAGAGAAATTCACACA 1320
Db 1468 GAGTGTGGAGAAAGCTTCAGATCAGCTCGCTTGTGTTCATCAGAAATTCACACT 1527
QY 1321 GGGGAGAAACCCGTATTAATGTAAAGAGTGGGAGAGGCTTTCATTTATGATCGAGCTTC 1380
Db 1528 GAGAGAGAGCC-TTTGAATGGAGTCCGATGAGCAAGAAAGCTCCCTCAGAGTTACAGCCGA 1586
QY 1381 GTGAAACATGAGAGAAATTC 1400

```

```

Db 1587 GTTAATATCAGAGATCCCA 1606

RESULT 13
US-11-136-527-3206
; Sequence 3206, Application US/11136527
; Publication No. US20050287570A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Mounts, William M
; TITLE OF INVENTION: Probe Arrays For Expression Profiling of Rat Genes
; FILE REFERENCE: 031896-041000 (AM101086)
; CURRENT APPLICATION NUMBER: US/11/136,527
; PRIOR FILING DATE: 2005-05-25
; PRIOR APPLICATION NUMBER: US 60/574,294
; NUMBER OF SEQ ID NOS: 362830
; SOFTWARE: Patencin version 3.2
; SEQ ID NO 3206
; LENGTH: 3350
; TYPE: DNA
; ORGANISM: Rattus norvegicus
US-11-136-527-3206

Query Match 19.9%; Score 401.4; DB 8; Length 3350;
Best Local Similarity 61.8%; Pred. No. 4,1e-117;
Matches 624; Conservative 7; Mismatches 378; Indels 0; Gaps 0;

QY 564 TAAAGAAATTCCTTTGAATGTAAAGACTGTGGGAGAGGCTTTAGTCGTGGCTATCACT 623
Db 1215 TGAGAGAAACGCTTGTGAATGTAAAGTGTAGTGTGCAAGTCTTCAGCTGAGCTCCCATCT 1274
QY 624 TAGTCAACATCAGAAATCCATATCTGTGAGAGAAACCTTATGATGTAAAGATGTAAAGA 683
Db 1275 CATTGCTACCAAGAACCATATACGGGAGAAACCGTACAAATGCAATCTGTGTGGGAA 1334
QY 684 GGCCTTCGTTGGGAGCAATCAGCTTATCAATCAATCAAAATTCATCTGGGAGAGAGCC 743
Db 1335 GTTCTTCACCCGAGCTCAAGTGTTCATCAGAGAGATTCATCTGAGAGAGAAACC 1394
QY 744 CTACGAATGTAAAGACTGTGGGAGAGGCTTTTGAATGGGCTCAAGCTCGTTATTCATTA 803
Db 1395 ATACAGGTGCATCTGTGTGGGAGAGCTTTCACCAAGATGTGTCTGTGGTGCATCA 1454
QY 804 GAGGATTCATCTGTGAGAAACCCATGATGTAAAGACTGTGAGAAAGGCTTTCGCGG 863
Db 1455 AAGAACTCACACTGGGAGAGGCTTGTATGATGCAACAGTGTGGAGTCTGTCCGCA 1514
QY 864 TGGTATGAGCTCACTCAGACCAAGAGATTCACACTGGGAGAGAAAGACTACGAATGCA 923
Db 1515 GAGCTTACAAACCTCATTTGCGCATCAAGAACCCACAGAGAAKAGCCGTACGAATGCAC 1574
QY 924 AGACTGTGGGAGAGCTTTAGCCGTGTGTATTAACCTTATTCAGCACAAAGAAATTCATAG 983
Db 1575 CCAGTGTGGGAGAGTCAATTCACAGATTCACAGCTCAGCGCATCAAGAAATTCACCTC 1634
QY 984 TGGGAGAGAGCCCTTCAGATGTAAAGACTGTGGGAGAGGCTTTATTTGTGTTCAAGCTT 1043
Db 1635 TGGAGAAACCTTATGATGTGAGTCACTGTGGCAAGTCTTCAGCCAAAGCTATTAAGCT 1694
QY 1044 CATTCAGCATTAAGAAATTCACACAGTGTGAAGAACCTTATGATGTCAAGAAATGTGGGAA 1103
Db 1695 TGTGTCCATTCAGAGAGCTCACACAKGAGAGAAACCTTCAGGTGTATCTAGTGGGAA 1754
QY 1104 GGCCTTATCTCGAGTCAATTAACCTTATCAGATCAGAGATTCACACCGGTGAGAGGCC 1163
Db 1755 ATCATTCAGCTGAGCTCMACAGCTGTGTCTCATCAAGAAACCCACAGCGKAGAGAGCC 1814
QY 1164 TCACGAATGTAAAGAGTGTGGGAGAGGCTTTTGGCTGGGGTTCGAGGCTGTAAAGACGA 1223
Db 1815 TTACKAGTGTAAAGAGTGTGGGAAATCATTAATCAATGCAAGCTCACACTTGTATGACCA 1874

```


US-10-750-185-30010

Query Match 19.3%; Score 388.8; DB 7; Length 1330;
Best Local Similarity 62.0%; Pred. No. 2.3e-113;
Matches 615; Conservative 0; Mismatches 377; Indels 0; Gaps 0;

Search completed: February 6, 2006, 22:13:57
Job time : 399 secs

```
QY 566 AGAGAAATTCCTTGAATGAAGACTGTGGGAAGCCCTTGTAGTCGTGCTATCACTTA 625
DB 992 AAGAGAAATGATTAATGTAAGATGTGAAGGTTTGTCCGCAAGGCTCATTTCA 933
QY 626 GTCAACATCAAAAATCCATCTGTGTGAAGAACTTATGAATGTAAAGATGTAGAAG 685
DB 932 TTCAACATCAAAAGGCTCACTGAGAGAAAACCTTCCAATGTATGAATGTGGAAAA 873
QY 686 CCTCCGTTGGGGCAATCAGCTTACTCAATCAAAAAATTCACTAGGGGAGAAAGCCCT 745
DB 872 GTTTTAGTCGAGTTCAATTTGTTATGAAATGAGAAATTCACACCGGGAGAAAGCCCT 813
QY 746 ACGAAATGTAAAGACTGTGGAAAGCTTTTGATGGGCTCAAGCCTCGTTATTCATPAGA 805
DB 812 ATGAGGTAAATTAATGCGAAAAACCTTATGTGTGAGCTCAACCTTATTAACATCAGA 753
QY 806 GGATTCATCTGTGTAAGAAACCTATGAATGTAAAGACTGTGAAAGGCTTCCGCGTG 865
DB 752 GAATCCACACTGGAGAGAGACCTATCAGTATCATCTGTAAACAGAGCTTCAGCCAGC 693
QY 866 GTGATGAGCTCACTGAGCAACAGAGATTCACACTGGGGGAAAGACTA CGAATGCCAAG 925
DB 692 GAAAGAACCTTGTAAACATCAGAGATTCACAGGTGAGAAACCCATTAATGTATGTG 633
QY 926 ACTGTGGGAAGACCTTTAGCCGTGTATTAATTCAGCACAAGAGAAATTCATATGTG 985
DB 632 ACTGTGGGAAGCTTTCAGTTGGAAGTCAACCTGATTTGAACATCAGAGAACACACTG 573
QY 986 GGGAGAAAGCTTACGAGTGAAGACTGTGGGAAGGCTTTTATTGTGTTCAAGCTCA 1045
DB 572 GTGAGAAACCTATCACTGTACCAATGTAGAAAAAGCTTTAGTAAATTCCTTACTTG 513
QY 1046 TTCAGATTAAGAAATTCACAGAGTGAAGAAACCTATGAATGTCAAGAAATGTGGAAAG 1105
DB 512 TTGAACATCAGAGAAATTCACACTGGGAAAGACCCATTAATGTGTGAATGTGGAAAG 453
QY 1106 CCTTTA CTGAGTCAATTAACCTTACTCAGCATCAGAAATCCACACGCTGAGAAAGCTTC 1165
DB 452 CCTTTAGATTAAGTACATATCTTATACAAACCAAAAAATCCACACGCTGAGAAAGCTTC 393
QY 1166 ACGAATGTAAAGAGTGTGGGAAGGCTTTGCTGGGCTTGAAGCTTCGTTAACAAGAGA 1225
DB 392 TTCTTTGTATGTGAATGTGAAAAAGCTTCAGTGGAGCTCATTCCTTATTGAACATCAGA 333
QY 1226 GGATACATACGGGCGAAGACCGTACAAAGTGA CAGAAATGTGGAAAGGCTTCATATTGTG 1285
DB 332 GGATTCATACAGGTGAAGACCTTATCAGTGCAGAAAGGTGGAAAAAGCTTCAGTCAAC 273
QY 1286 GCTATCACTCACTCAGCAGAGAGAAATCCACA CAGGCGAAACCCGTTATTAATGTAAAG 1345
DB 272 TTTCGAACCTTACTCGCATCAGAGAAATTCACA CAGGCGCAAAACCCATTAATGTGAAG 213
QY 1346 AGTGTGGAAAGCTTTCATTATGATCGAGCTCGTGAACATGAGAGAAATTCATACCG 1405
DB 212 AATGTGAAAAAGCCTTTAGTGAAGCTCAGGTCTTATTCAGCACAGAGAAATCCACACCA 153
QY 1406 GGGTGAACCTTANNGGTGACAGAAATGTGGGAAGGCTTATGTCACGCGCATCAGCTTA 1465
DB 152 GAGAGAAAGACTTATTCATACATGAAGAACCAAGAAAGTTTGAACCAATGAGAGTCTTG 93
QY 1466 CACAAATCAAGAAAACGACAGTGGGCGAAATCTTACGAATGTAAAGAGTGGCGAAG 1525
DB 92 TTATACACAGGAAGTACCTACCTTAAGGAAAAATCTTATAATGTGAATGTGGGAAA 33
QY 1526 CATGTAAACCACTAAACATCTCGAGAACAT 1557
DB 32 CTTTCAGTGTCAAGTGTCAATCTTGTGCACAT 1
```


GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceeleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 6, 2006, 21:33:22 ; Search time 1666 Seconds
(without alignments)
10001.679 Million cell updates/sec

Title: US-10-111-257-3

Perfect score: 2015
Sequence: 1 cgcctcgacacgcgtgacgc.....gtgaagttccacttaaa 2015

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 9793542 seqs, 4134689005 residues

Total number of hits satisfying chosen parameters: 19587084

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA Main:*

- 1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
- 2: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq.*
- 3: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq.*
- 4: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq.*
- 5: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq.*
- 6: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq.*
- 7: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq.*
- 8: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq.*
- 9: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq.*
- 10: /cgn2_6/ptodata/1/pubpna/US11_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1967.4	97.6	2173	7	US-10-221-625-108 Sequence 108, App
2	1946.2	96.6	3293	3	US-09-764-864-123 Sequence 123, App
3	1420.2	70.5	1453	7	US-10-296-115-424 Sequence 424, App
4	1397.8	69.4	1429	7	US-10-641-643-689 Sequence 689, App
5	734.6	36.5	4483	9	US-10-499-353A-332 Sequence 332, App
6	703	34.9	2143	3	US-09-881-578-3 Sequence 3, Appl1
7	703	34.9	2143	3	US-10-623-914-3 Sequence 3, Appl1
8	618	30.7	1945	6	US-10-094-749-1120 Sequence 1120, Ap
9	615.4	30.5	2235	6	US-10-108-260A-2359 Sequence 2359, A
10	603.6	30.0	1900	3	US-10-450-763-25299 Sequence 25299, A
11	593.6	29.5	774	3	US-09-764-864-567 Sequence 567, App
12	584.4	29.0	592	6	US-10-029-386-13162 Sequence 13162, A
13	546	26.7	3163	9	US-10-786-148-116 Sequence 116, App
14	538.6	26.1	1813	7	US-10-121-625-162 Sequence 162, App
15	538	26.7	2334	6	US-10-094-749-1085 Sequence 1085, Ap
16	538	26.7	4683	9	US-10-450-763-5630 Sequence 5630, Ap
17	520.6	25.8	2274	5	US-10-037-270-154 Sequence 154, App
18	520.6	25.8	2274	6	US-10-117-722-154 Sequence 154, App
19	520.6	25.3	526	3	US-09-954-531-545 Sequence 545, App
20	510.2	25.3	526	3	US-10-843-641A-1612 Sequence 1612, Ap
21	510.2	25.3	2499	6	US-10-120-988-331 Sequence 331, App
22	510	25.3	2499	6	US-10-450-763-12507 Sequence 12507, A
23	506.6	25.1	4216	9	US-10-450-763-12507 Sequence 12507, A

24	506.6	25.1	8907	9	US-10-450-763-3997 Sequence 3997, Ap
25	503.4	25.0	1635	3	US-09-864-761-16962 Sequence 16962, A
26	502.8	25.0	2075	5	US-10-098-841-110 Sequence 110, App
27	502.8	25.0	3472	6	US-10-104-047-415 Sequence 415, App
28	500.8	24.9	1931	6	US-10-108-260A-113 Sequence 113, App
29	500.8	24.9	2521	6	US-10-094-749-1398 Sequence 1398, Ap
30	499	24.8	2266	6	US-10-094-749-418 Sequence 418, App
31	495.6	24.6	1268	6	US-10-029-386-25225 Sequence 25225, A
32	495.4	24.6	1694	6	US-10-104-047-1500 Sequence 1500, Ap
33	489.6	24.3	1907	7	US-10-121-625-142 Sequence 142, App
34	488.8	24.3	3139	5	US-10-198-846-13470 Sequence 13470, A
35	488.2	24.2	990	3	US-09-864-761-26905 Sequence 26905, A
36	486.8	24.2	1919	3	US-09-864-761-10504 Sequence 10504, A
37	486	24.1	4797	5	US-10-171-311-231 Sequence 231, App
38	486	24.1	4797	8	US-10-645-756-43 Sequence 43, Appl
39	483.8	24.0	1716	6	US-10-029-386-24986 Sequence 24986, A
40	478.2	23.7	2925	5	US-10-037-270-163 Sequence 163, App
41	478.2	23.7	2925	6	US-10-117-722-163 Sequence 163, App
42	478.2	23.7	2925	9	US-10-122-851-163 Sequence 163, App
43	474.8	23.6	2241	6	US-10-104-047-693 Sequence 693, App
44	473.2	23.5	5830	6	US-10-242-355-960 Sequence 960, App
45	464.4	23.0	4132	5	US-10-198-846-13133 Sequence 13133, A

ALIGNMENTS

RESULT 1
US-10-221-625-108
Sequence 108, Application US/10221625
Publication No. US2004003942A1
GENERAL INFORMATION:
APPLICANT: INCYTE GENOMICS, INC.
APPLICANT: HILTMAN, Jennifer L.
APPLICANT: BAUGHN, Mariah R.
APPLICANT: YUE, Henry
APPLICANT: LAL, Preeti
APPLICANT: LU, Dyung Aina M.
APPLICANT: PATTERSON, Chandra
APPLICANT: AZIMZAI, Yalda
APPLICANT: BANDMAN, Olga
APPLICANT: TANG, Y. Tom
APPLICANT: MATHUR, Preeti
APPLICANT: SHAH, Purvi
APPLICANT: AU-YOUNG, Janice
APPLICANT: REDDY, Roopa
TITLE OF INVENTION: TRANSCRIPTION FACTORS
FILE REFERENCE: PF-0761 PCT
CURRENT APPLICATION NUMBER: US/10/221,625
CURRENT FILING DATE: 2001-03-13
NUMBER OF SEQ ID NOS: 214
SOFTWARE: PERL Program
SEQ ID NO 108
LENGTH: 2173
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
OTHER INFORMATION: Incyte ID No. US2004003942A1 095210CB1
US-10-221-625-108
Query Match 97.6%; Score 1967.4; DB 7; Length 2173;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1968; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY CAGGCACGATCCTTCAGAAAAAGCATCCCGAGAGAGACGATGTTAAACATCTT 106
DB CAGGCACGATCCTTCAGAAAAAGCATCCCGAGAGAGACGATGTTAAACATCTT 248
QY AGGTGAGCTTAGCCTTCGGAATTTGCTTCTTCAGTGAAGACCCGAGAACTGATC 166
DB AGGTGAGCTTAGCCTTCGGAATTTGCTTCTTCAGTGAAGACCCGAGAACTGATC 308

```

OY 167 AGTCTTCAGTCTTAATAAATGAGCCAGAGGTTGGTGAAGTTCGCGAGATGACCATAG 226
DB 309 AGTTCTTCAGTCTTAATAAATGAGCCAGAGGTTGGTGAAGTTCGCGAGATGACCATAG 368
OY 227 ACTTTTTCAGAGAGAGTGGGCTGTCTAACTCTGCTCAAGAGGACCTGTACTGGAGAG 286
DB 369 ACTTTTTCAGAGAGAGTGGGCTGTCTAACTCTGCTCAAGAGGACCTGTACTGGAGAG 428
OY 287 TGATGCTGGAGAACTACAGTAACTTGGTCTCACTGGATTGGAGTCAAGATTAAGAAATA 346
DB 429 TGATGCTGGAGAACTACAGTAACTTGGTCTCACTGGATTGGAGTCAAGATTAAGAAATA 488
OY 347 AGAGTTTACCTACAGAAAAAAACATTCATGAATAAAGGCTTCCAAAAGAAATTCAGATA 406
DB 489 AGAGTTTACCTACAGAAAAAAACATTCATGAATAAAGGCTTCCAAAAGAAATTCAGATA 548
OY 407 GAAAGAAATTAATCCTTGGCCGTAATGGAATATGTAAGTACGCTTGAAGAACCACAGC 466
DB 549 GAAAGAAATTAATCCTTGGCCGTAATGGAATATGTAAGTACGCTTGAAGAACCACAGC 608
OY 467 GCTCCAGAGGAGGTATGCAATCAGATGATCATCAATTATGCAAAAAGGCTGCTACTA 526
DB 609 GCTCCAGAGGAGGTATGCAATCAGATGATCATCAATTATGCAAAAAGGCTGCTACTA 668
OY 527 GAGAAAGGACCCCTCTAGAACACATCAGAGACATCAAGAAGAAATTCCTTTGAATGTA 586
DB 669 GAGAAAGGACCCCTCTAGAACACATCAGAGACATCAAGAAGAAATTCCTTTGAATGTA 728
OY 587 AGGACTGTGGGAAAGGCTTTAAGTGTGAGTATCACTTAAGTCAAGAAATTCAGATA 646
DB 729 AGGACTGTGGGAAAGGCTTTAAGTGTGAGTATCACTTAAGTCAAGAAATTCAGATA 788
OY 647 CTGGTGAGAAACCTTAATGATGTAAGAAATGTAAGAGGCTTCGCTGGGAGCAATCAGC 706
DB 789 CTGGTGAGAAACCTTAATGATGTAAGAAATGTAAGAGGCTTCGCTGGGAGCAATCAGC 848
OY 707 TTAATCAACATCAAAAAATTCATACTGAGGAGAGGCTTACGAATGAAGAATCTGGGA 766
DB 849 TTAATCAACATCAAAAAATTCATACTGAGGAGAGGCTTACGAATGAAGAATCTGGGA 908
OY 767 AGGCTTTTGAATGGGCTCAAGCTCTGTTATTAATGAAGATTCATCTGGTGAAGAAAC 826
DB 909 AGGCTTTTGAATGGGCTCAAGCTCTGTTATTAATGAAGATTCATCTGGTGAAGAAAC 968
OY 827 CCTATGATGTAAGAGCTGTGAAAGGCTTCGCGTGGTGAATGAGTCACTCAGACAC 886
DB 969 CCTATGATGTAAGAGCTGTGAAAGGCTTCGCGTGGTGAATGAGTCACTCAGACAC 1028
OY 887 AGAGATTCACACTGGGAGAGAAAGACTACGAATGCAAGAATCTGGGAGAGACCTTTAGCC 946
DB 1029 AGAGATTCACACTGGGAGAGAAAGACTACGAATGCAAGAATCTGGGAGAGACCTTTAGCC 1088
OY 947 GTGTGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1006
DB 1089 GTGTGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1148
OY 1007 AAGAAGCTGGGAGAGGCTTTAATTTGTGCTCAAGCTCATTCAGCATTAATAAAGATTCACA 1066
DB 1149 AAGAAGCTGGGAGAGGCTTTAATTTGTGCTCAAGCTCATTCAGCATTAATAAAGATTCACA 1208
OY 1067 CAGGTGAGAAACCTTAATGAATGTCAAGAAATGTGGAGAGGCTTTACTCGAGTCAATTAAC 1126
DB 1209 CAGGTGAGAAACCTTAATGAATGTCAAGAAATGTGGAGAGGCTTTACTCGAGTCAATTAAC 1268
OY 1127 TTAATCAAGATCAAGAAATTCACACCGGTGAGAAAGCTTCAGAAATGAAGATGTGGGA 1186
DB 1269 TTAATCAAGATCAAGAAATTCACACCGGTGAGAAAGCTTCAGAAATGAAGATGTGGGA 1328
OY 1187 AGGCTTTTGGCTGGGCTGAGGCTGTGAAGACGAGAGATACATACGGGCGAGAGAGC 1246
DB 1329 AGGCTTTTGGCTGGGCTGAGGCTGTGAAGACGAGAGATACATACGGGCGAGAGAGC 1388
OY 1247 CGTACAAGTGACAGAAATGTGGAGAGGCTTCAATTTGTGCTATCACCTCAGACAGC 1306

```

```

DB 1389 CGTACAAGTGACAGAAATGTGGAGAGGCTTCAATTTGGCTATACCTCAGTCAAGCAG 1448
OY 1307 AGAGAAATCAACACAGAGCAAAACCCGTAATTAATGTAAGAGTGGAGAGCTTTCATTT 1366
DB 1449 AGAGAAATCAACACAGAGCAAAACCCGTAATTAATGTAAGAGTGGAGAGCTTTCATTT 1508
OY 1367 ATGATTCAGAGCTCTGTGAACATGAGAGAAATTCATACCGGGGTGAACCTTAATGGGTGA 1426
DB 1509 ATGATTCAGAGCTCTGTGAACATGAGAGAAATTCATACCGGGGTGAACCTTAATGGGTGA 1568
OY 1427 CAGAAATGTGGAGAGGCTTTAATGTCACGGCCATCAGCTTACACAAATCAGAAAAAGCACA 1486
DB 1569 CAGAAATGTGGAGAGGCTTTAATGTCACGGCCATCAGCTTACACAAATCAGAAAAAGCACA 1628
OY 1487 GTGGGGCGAAATCCTACGAATGTAAGAGTGGGAGAGGATGTAAACCACTTAACATC 1546
DB 1629 GTGGGGCGAAATCCTACGAATGTAAGAGTGGGAGAGGATGTAAACCACTTAACATC 1688
OY 1547 TCCGAGAAACATCAGAGATCCACAAACAGTTGAAGAGCTTTTGAACGAGTACCGGCTC 1606
DB 1689 TCCGAGAAACATCAGAGATCCACAAACAGTTGAAGAGCTTTTGAACGAGTACCGGCTC 1748
OY 1607 GTATCTATGCTTTCGCTTTCACAGTTTGTATCTGCAATGCAATGCAATTAATAAT 1666
DB 1749 GTATCTATGCTTTCGCTTTCACAGTTTGTATCTGCAATGCAATGCAATTAATAAT 1808
OY 1667 TAAATGAGAAATTTCCAGAAATTAAGAAATTTAAGTCTCAAAATGAGTGGCTTTCAGATA 1726
DB 1809 TAAATGAGAAATTTCCAGAAATTAAGAAATTTAAGTCTCAAAATGAGTGGCTTTCAGATA 1868
OY 1727 GCGTGAATGAATCTCTGCTGTCGGGCTCCAGCCGCGGGAGTGAATCAATCCCTTGG 1786
DB 1869 GCGTGAATGAATCTCTGCTGTCGGGCTCCAGCCGCGGGAGTGAATCAATCCCTTGG 1928
OY 1787 TCCAGACATCAACAGCTGTATACGCCACCCAGCTTGTATGTAAGTGAAGGCTTTTGG 1846
DB 1929 TCCAGACATCAACAGCTGTATACGCCACCCAGCTTGTATGTAAGTGAAGGCTTTTGG 1988
OY 1847 TGATCAGATCAATTAATCCAGAGTCAAGTGGCTGTGCCAAGTCACTCACTTGTGCTT 1906
DB 1989 TGATCAGATCAATTAATCCAGAGTCAAGTGGCTGTGCCAAGTCACTCACTTGTGCTT 2048
OY 1907 AACAGTGGCCCGAGAGAGAGAGTATGATGATCTGGTGAATGCAATGCAAAAGAGAG 1966
DB 2049 AACAGTGGCCCGAGAGAGAGAGTATGATGATCTGGTGAATGCAATGCAAAAGAGAG 2108
OY 1967 CCAAGAGTGTCTCTTTTAATGAAAGGTGAAGTTCACACTTAA 2015
DB 2109 CCAAGAGTGTCTCTTTTAATGAAAGGTGAAGTTCACACTTAA 2157

```

RESULT 2

```

US-09-764-864-123
; Sequence 123, Application US/09764864
; Patent No. US20020132753A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT223
; CURRENT APPLICATION NUMBER: US/09/764,864
; PRIORITY FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1792
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 123
; LENGTH: 3293
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-864-123

```

```

Query Match      96.6%; Score 1946.2; DB 3; Length 3293;
Best Local Similarity 98.3%; Pired. No. 0;

```


Publication No. US20040053248A1
 GENERAL INFORMATION:
 APPLICANT: Hyseq Inc
 TITLE OF INVENTION: No. US20040053248A1el Nucleic Acids and Polypeptides
 FILE REFERENCE: 784FCT
 CURRENT APPLICATION NUMBER: US/10/296,115
 CURRENT FILING DATE: 2002-11-18
 PRIOR APPLICATION NUMBER: US09/488,725
 PRIOR FILING DATE: 2000-01-21
 PRIOR APPLICATION NUMBER: US09/552,317
 PRIOR FILING DATE: 2000-04-25
 NUMBER OF SEQ ID NOS: 1478
 SEQ ID NO 424
 LENGTH: 1453
 TYPE: DNA
 ORGANISM: Homo sapiens
 US-10-296-115-424

Query Match 70.5%; Score 1420.2; DB 7; Length 1453;
 Best Local Similarity 99.5%; Pred. No. 0;
 Matches 1446; Conservative 0; Mismatches 3; Indels 4; Gaps 2;

QY 567 GGAGAAATTCCTTGAATGTAAGAGCTGTGGAAAGCCTTTAGTCGTGCTATCAACTTGA 626
 DB 1453 GGAGAAATTCCTTGAATGTAAGAGCTGTGGAAAGCCTTTAGTCGTGCTATCAACTTGA 1394
 QY 627 TCAACATCAGAAATCCATCTACTGTGAGAAACCTTATGTAATGTAAGAAATGTAAGAGC 686
 DB 1393 TCAACATCAGAAATCCATCTACTGTGAGAAACCTTATGTAATGTAAGAAATGTAAGAGC 1334
 QY 687 CTTCCTGGGGCAATCACTTACTCAACATCAAAATTTATCTGGGGAGAAAGCCCTTA 746
 DB 1333 CTTCCTGGGGCAATCACTTACTCAACATCAAAATTTATCTGGGGAGAAAGCCCTTA 1274
 QY 747 CGAATGTAAGAGCTGTGGAAAGCCTTTGATGGGCTCAAGCCTGTTATCATTAAGAG 806
 DB 1273 CGAATGTAAGAGCTGTGGAAAGCCTTTGATGGGCTCAAGCCTGTTATCATTAAGAG 1214
 QY 807 GATTCACTGTGTGAAGAAACCTATGTAATGTAAGAGCTGTGAAGAGCCTTTGGGCTGG 866
 DB 1213 GATTCACTGTGTGAAGAAACCTATGTAATGTAAGAGCTGTGAAGAGCCTTTGGGCTGG 1154
 QY 867 TGATGAGCTCACTCAGACCAAGAGATTCACACTGGGGAGAAAGACTTACGAATGCAAGA 926
 DB 1153 TGATGAGCTCACTCAGACCAAGAGATTCACACTGGGGAGAAAGACTTACGAATGCAAGA 1094
 QY 927 CTGTGGGAAGACCTTTAGCCGTGTATTAACCTTATTCAGACCAAGAGATTCATAGTGG 986
 DB 1093 CTGTGGGAAGACCTTTAGCCGTGTATTAACCTTATTCAGACCAAGAGATTCATAGTGG 1034
 QY 987 GGAGAAAGCTTACGAGTGTAAAGAGCTGTGGAAAGCCTTTATTTGTGTCAAGCCTCAT 1046
 DB 1033 GGAGAAAGCTTACGAGTGTAAAGAGCTGTGGAAAGCCTTTATTTGTGTCAAGCCTCAT 974
 QY 1047 TCAGCATAAAGAAATTCACACAGGTGAGAAACCTTATGAAATGTCAGAAATGTGGAAAGC 1106
 DB 973 TCAGCATAAAGAAATTCACACAGGTGAGAAACCTTATGAAATGTCAGAAATGTGGAAAGC 914
 QY 1107 CTTTACTCGAGTCAATTAACCTTACTCAGCTCAGAAATTCACACCGGTGAGAAAGCCTCA 1166
 DB 913 CTTTACTCGAGTCAATTAACCTTACTCAGCTCAGAAATTCACACCGGTGAGAAAGCCTCA 854
 QY 1167 CGAATGTAAGAGAGTGTGGAAAGCCTTTCCCTGGGGTTCAGAGCCTGTTAAGACAGAG 1226
 DB 853 CGAATGTAAGAGAGTGTGGAAAGCCTTTCCCTGGGGTTCAGAGCCTGTTAAGACAGAG 794
 QY 1227 GATACATACGGGCGAGAAAGCCTTACAAAGTGCACAGAAATGTGGAAAGCCTTCAATTGTGG 1286
 DB 793 GATACATACGGGCGAGAAAGCCTTACAAAGTGCACAGAAATGTGGAAAGCCTTCAATTGTGG 724
 QY 1287 CTATCACCCTCAGCAGACGAGAAATTCACACAGGGGAAACCCCGTATTAATGTAAAGA 1346
 DB 733 CTATCACCCTCAGCAGACGAGAAATTCACACAGGGGAAACCCCGTATTAATGTAAAGA 674

QY 1347 GTGTGGAAAGCCTTCAATTAATGATCGAGCCTGTGAAACATGAGAAATTCATACCGG 1406
 DB 673 GTGTGGAAAGCCTTCAATTAATGATCGAGCCTGTGAAACATGAGAAATTCATACCGG 614
 QY 1407 GGTGAAACCTTATGGGTGTACGAATGTGGAAAGCCTTTACTATCAGGCTATCAGCTTAC 1466
 DB 613 GGTGAAACCTTATGGGTGTACGAATGTGGAAAGCCTTTACTATCAGGCTATCAGCTTAC 554
 QY 1467 ACAACATCAGAAAGAGCAGAGGGGCGAAATCCTACGAATGTAAGAGTGGGGAAGGC 1526
 DB 553 ACAACATCAGAAAGAGCAGAGGGGCGAAATCCTACGAATGTAAGAGTGGGGAAGGC 494
 QY 1527 ATGTAAACCACTTAAACATCTCCAGAAATCAGAGAGATTCACAAAGTTGAAGAGCCTT 1586
 DB 493 ATGTAAACCACTTAAACATCTCCAGAAATCAGAGAGATTCACAAAGTTGAAGAGCCTT 434
 QY 1587 TTGAACGAGTAGCCCGCTCGTATCTATGTTGCTTCCACAGTTGTTACCTGCAAT 1646
 DB 433 TTGAACGAGTAGCCCGCTCGTATCTATGTTGCTTCCACAGTTGTTACCTGCAAT 374
 QY 1647 CAACGAGTTCAGAAATATTAATGAAATTCAGAAATTAAGAAATTTAAGTCTCAA 1706
 DB 373 CAACGAGTTCAGAAATATTAATGAAATTCAGAAATTAAGAAATTTAAGTCTCAA 314
 QY 1707 ATGTGTGCTTCTGAGTAGCGGTGATGAATCTCTGCTGTCCGCTCCAGCCGCGCG 1766
 DB 313 ATGTGTGCTTCTGAGTAGCGGTGATGAATCTCTGCTGTCCGCTCCAGCCGCGCG 254
 QY 1767 GATGTGATCATCTCTGTCAGACATCCACGCTGTATACGCCACCCACCTGCTGAG 1826
 DB 253 GATGTGATCATCTCTGTCAGACATCCACGCTGTATACGCCACCCACCTGCTGAG 194
 QY 1827 TGACTTAGTAGCCGCTTGTGTATGATCAATCACTATCCAGCA--TGACAGTCCGTGCGC 1885
 DB 193 TGACTTAGTAGCCGCTTGTGTATGATCAATCACTATCCAGCAATTCAGTCTGTGAC 134
 QY 1886 CAAGTAGTCTCACTTGTCTTAAAGTGGCCCAAGAGCAGAGTAG--TGATCTGG 1942
 DB 133 CAAGTAGTCTCACTTGTCTTAAAGTGGCCCAAGAGCAGAGTAGATGATCTGG 74
 QY 1943 TGATTTGATATGCCAAGAGAGCCACAAAGTGTCTCTTTTAAATGAAAGGTGAAAG 2002
 DB 73 TGATTTGATATGCCAAGAGAGCCACAAAGTGTCTCTTTTAAATGAAAGGTGAAAG 14
 QY 2003 TTCTCACTTAA 2015
 DB 13 TTCTCACTTAA 1

RESULT 4
 US-10-641-643-689
 Sequence 689, Application US/10641643
 Publication No. US20040077003A1
 GENERAL INFORMATION:
 APPLICANT: Cocks, Benjamin G.
 Susan G. Stuart
 Jeffrey J. Seilhamer
 TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL
 GENE EXPRESSION
 NUMBER OF SEQUENCES: 1508
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
 STREET: 3174 PORTER DRIVE
 CITY: PALO ALTO
 STATE: CALIFORNIA
 COUNTRY: USA
 ZIP: 94304
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/641,643
FILING DATE: 14-Aug-2003
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: <Unknown>
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Zeller, Karen J.
REGISTRATION NUMBER: 37,071
REFERENCE/DOCKET NUMBER: PA-0001 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650) 855-0555
TELEFAX: (650) 845-4166
INFORMATION FOR SEQ ID NO: 689:
SEQUENCE CHARACTERISTICS:
LENGTH: 1429 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: TLYMN02
CLONE: 450088
SEQUENCE DESCRIPTION: SEQ ID NO: 689 :
US-10-641-643-689
Query Match 69.4%; Score 1397.8; DB 7; Length 1429;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 1411; Conservative 0; Mismatches 7; Indels 2; Gaps 1;
QY 555 GGGAGGCGTTAGTCGTCGCTATCACTTACTCAACATGAGAAATCCATCTGTGTG 654
DB 10 GGGAGGCGTTAGTCGTCGCTATCACTTACTCAACATGAGAAATCCATCTGTGTG 69
QY 655 AAACCTTAGATGAATGAATGAATGAATGAATGAATGAATGAATGAATGAATGAAT 714
DB 70 AAACCTTAGATGAATGAATGAATGAATGAATGAATGAATGAATGAATGAATGAAT 129
QY 715 CATCAAAAATTCATCTGCGGAGAAAGCCCTACGAATGTAAGACTGTGGGAAAGGCTT 772
DB 130 CATCAAAAATTCATCTGCGGAGAAAGCCCTACGAATGTAAGACTGTGGGAAAGGCTT 189
QY 773 TTGCGATGGGCTCAAGCCCTGTTATTCATGAAGATTTATCTGCTGTAAGAAACCTT 832
DB 190 TTGCGATGGGCTCAAGCCCTGTTATTCATGAAGATTTATCTGCTGTAAGAAACCTT 249
QY 833 AATGTAAGACTGTGGAAGGCTTTCGCGTGTGATGAGCTCACTGACACAGAGAT 892
DB 250 AATGTAAGACTGTGGAAGGCTTTCGCGTGTGATGAGCTCACTGACACAGAGAT 309
QY 893 TTCACACTGCGGAGAAAGACTACGAATGCAAGACTGTGGGAAAGCTTTAGCCGTGT 952
DB 310 TTCACACTGCGGAGAAAGACTACGAATGCAAGACTGTGGGAAAGCTTTAGCCGTGT 369
QY 953 ATAACTTATTCAGACAGAAATTCATGTCGCGGAGAAAGCTTACGATGTAAGACT 1012
DB 370 ATAACTTATTCAGACAGAAATTCATGTCGCGGAGAAAGCTTACGATGTAAGACT 429
QY 1013 GTGGAAGGCTTTTATTTGTGTGTTCAAGCCCTATTCAGATTAAGAAATTCACAGGTG 1072
DB 430 GTGGAAGGCTTTTATTTGTGTGTTCAAGCCCTATTCAGATTAAGAAATTCACAGGTG 489
QY 1073 AGAAACCTTATGATGTCAGAAATGTGGGAGGCTTTTACGAGTCAATTACCTTACTC 1132
DB 490 AGAAACCTTATGATGTCAGAAATGTGGGAGGCTTTTACGAGTCAATTACCTTACTC 549
QY 1133 AGCATCAGAAAGATCCACACCGGTGGAAGCTTCAGAAATGTAAGAGGTGGAGGCTT 1192
DB 550 AGCATCAGAAAGATCCACACCGGTGGAAGCTTCAGAAATGTAAGAGGTGGAGGCTT 609
QY 1193 TTGCGTGGGCTTCGAGCCTGTTAAGCAAGAGGATACATACGCGCGAAGCCGTACA 1252
DB 610 TTGCGTGGGCTTCGAGCCTGTTAAGCAAGAGGATACATACGCGCGAAGCCGTACA 669

QY 1253 AGTGCACGAATGTGGGAAGGCTTCAATTGTGCTATCACTCACTGACAGACAGAGA 1312
DB 670 AGTGCACGAATGTGGGAAGGCTTCAATTGTGCTATCACTCACTGACAGACAGAGA 729
QY 1313 TCACACAGGCGAAACCCCGTTAATGTAAGAGTGTGGGAAGGCTTCAATTATGAT 1372
DB 730 TCACACAGGCGAAACCCCGTTAATGTAAGAGTGTGGGAAGGCTTCAATTATGAT 789
QY 1373 CGAGCCTGTAAGACATGAGAAATTCATCCGCGGTGAAACCTTATGCTATCAAGAT 1432
DB 790 CGAGCCTGTAAGACATGAGAAATTCATCCGCGGTGAAACCTTATGCTATCAAGAT 849
QY 1433 GTGGGAAGCTTTAGTCAAGGCTTACGCTTACCAACATGAGAAAGGCAAGTGGG 1492
DB 850 GTGGGAAGCTTTAGTCAAGGCTTACGCTTACCAACATGAGAAAGGCAAGTGGG 909
QY 1493 CGAAATCCTACGAATGTAAGAGTCCGGAAGGCAATGTAACCACTTAACATCTCCAG 1552
DB 910 CGAAATCCTACGAATGTAAGAGTCCGGAAGGCAATGTAACCACTTAACATCTCCAG 969
QY 1553 AACATCAGAGATCCACAAACAGTTGAGAGCCCTTTGAAACGAGTAGCCCGCTGTATCT 1612
DB 970 AACATCAGAGATCCACAAACAGTTGAGAGCCCTTTGAAACGAGTAGCCCGCTGTATCT 1029
QY 1613 ATGTTTGGCTTTCCACAGTTTGTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1672
DB 1030 ATGTTTGGCTTTCCACAGTTTGTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1089
QY 1673 GAAATTCAGAAATTAAGAAATTTAAGTCTCAAAATGAGTCCCTTGAAGAGGTGA 1732
DB 1090 GAAATTCAGAAATTAAGAAATTTAAGTCTCAAAATGAGTCCCTTGAAGAGGTGA 1149
QY 1733 TGAATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1792
DB 1150 TGAATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1209
QY 1793 ACATCCAGGCTGTATACGACCCACCTGCTGATGATTAAGTCCGCTTGTGTATCA 1852
DB 1210 ACATCCAGGCTGTATACGACCCACCTGCTGATGATTAAGTCCGCTTGTGTATCA 1269
QY 1853 GATCAACTATCCACATCAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1912
DB 1270 GATCAACTATCCACATCAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1329
QY 1913 GGGCCAGAGAGCAGAGTATGATGCTGTGATTCGATATGCCAAAGAGAGCCACAA 1972
DB 1330 GGGCCAGAGAGCAGAGTATGATGCTGTGATTCGATATGCCAAAGAGAGCCACAA 1389
QY 1973 AGTGCTTCTTTAATGAAAGGTGAAGTTCGACTT 2012
DB 1390 AGTGCTTCTTTAATGAAAGGTGAAGTTCGACTT 1429

RESULT 5
US-10-499-353A-332
; Sequence 332, Application US/10499353A
; Publication No. US20050123551A1
; GENERAL INFORMATION:
; APPLICANT: diadexus, Inc.
; APPLICANT: Sun, Yongming
; APPLICANT: Liu, Chenghua
; TITLE OF INVENTION: Compositions and Methods Relating to Endometrial Specific Genes ar
; FILE REFERENCE: DEX-0377
; CURRENT APPLICATION NUMBER: US/10/499,353A
; CURRENT FILING DATE: 2004-06-17
; PRIOR APPLICATION NUMBER: US 60/342,751
; PRIOR FILING DATE: 2001-12-21
; NUMBER OF SEQ ID NOS: 66
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 332
; LENGTH: 4483

TYPE: DNA
ORGANISM: Homo sapien
US-10-499-353A-332

Query Match 36.5%; Score 734.6; DB 9; Length 4483;
Best Local Similarity 91.2%; Pred. No. 2e-215;
Matches 791; Conservative 0; Mismatches 74; Indels 2; Gaps 1;

```
QY 47 CAGGCGAGCATCCTTCAGAAAAAGCATCCCGAGGAGGAAGCGATCGTTAAACATCTT 106
DB 1184 CAGGCGAGCATCCTTCAGAAAAAGCATCCCGAGGAGGAAGCGATCGTTAAACATCTT 1243
QY 107 AGGTGAGCTCTAGCCCTCTCGAATTTGCTTCTTCAGTGAAACCCCGAAGACTGATC 166
DB 1244 AGGTGAGCTCTAGCCCTCTCGAATTTGCTTCTTCAGTGAAACCCCGAAGACTGATC 1303
QY 167 AGTTCCTCAGTTCTTAAACATATGCGCCAGGGTTTGGTGACGTTCCCGACGTAAGCCATG 226
DB 1304 AGTTCCTCAGTTCTTAAACATATGCGCCAGGGTTTGGTGACGTTCCCGACGTAAGCCATG 1363
QY 227 ACTTTTCTCAGAGGAGTGGGCTGTCTGAACTCTGCTCAGAGGAGCTGTACTGGGAGC 286
DB 1364 ACTTTTCTCAGAGGAGTGGGCTGTCTGAACTCTGCTCAGAGGAGCTGTACTGGGAGC 1423
QY 287 TGATGCTGAGAACTACAGTAACCTTGCTCACTGGAATTTGAGTACGATATGAAATA 346
DB 1424 TGATGCTGAGAACTACAGTAACCTTGCTCACTGGAATTTGAGTACGATATGAAATA 1483
QY 347 AAGATTACCTACAGAAAAAACATTCATGAAATAAGGGCTTCAGAAAGAAATTCAGTA 406
DB 1484 AAGATTACCTACAGAAAAAACATTCATGAAATAAGGGCTTCAGAAAGAAATTCAGTA 1543
QY 407 GAAGAAATAATCCCTTGGCCGTACCTGATGTGATGTAAGGATGCTTGAAGACACAGC 466
DB 1544 GAAGAAATAATCCCTTGGCCGTACCTGATGTGATGTAAGGATGCTTGAAGACACAGC 1603
QY 467 GCTCCAGAGGAGTATGTCAATCAGATGATCAATTAATGTCAAAAGCCCTGCTACTA 526
DB 1604 GCTCCAGAGGAGTATGTCAATCAGATGATCAATTAATGTCAAAAGCCCTGCTACTA 1663
QY 527 GAGAAAGCACCCCTCCTAAGAACATCAAGACATCATTAAGGAATTCCTTTGAATGTA 586
DB 1664 GAGAAAGCACCCCTCCTAAGAACATCAAGACATCATTAAGGAATTCCTTTGAATGTA 1723
QY 587 AGGACTGTGGGAAGGCTTTAGTGTGCTGATCAACTTAAGTCAACATCAAGAAATCCATA 646
DB 1724 AGGACTGTGGGAAGGCTTTAGTGTGCTGATCAACTTAAGTCAACATCAAGAAATCCATA 1783
QY 647 CTGTGAGAAACCTTAATGTAATGTAAGAAATGTAAGAGCCCTTCGTTGGGCAATCAGC 706
DB 1784 CTGTGAGAAACCTTAATGTAATGTAAGAAATGTAAGAGCCCTTCGTTGGGCAATCAGC 1843
QY 707 TTAATCAATCAAAAAATTCATATCTGGGGAAGCCCTACGAATGTAAGACTGTGGGA 766
DB 1844 TTAATCAATCAAAAAATTCATATCTGGGGAAGCCCTACGAATGTAAGACTGTGGGA 1901
QY 767 AGGCTTTTGGATGGGCTCAAGCCGTGTTATTCATTAAGAGATTCAATACGTGTGAAAAAC 826
DB 1902 AGGCTTTTGGATGGGCTCAAGCCGTGTTATTCATTAAGAGATTCAATACGTGTGAAAAAC 1961
QY 827 CCTATGAAATGTAAGAATGTGGAAGGCTTTTGGCGCTGTGTGATGAGCTCACTCAGACCC 886
DB 1962 CCTATGAAATGTAAGAATGTGGAAGGCTTTTGGCGCTGTGTGATGAGCTCACTCAGACCC 2021
QY 887 AAGAGATTCCACACTGGGAGGAAGAACT 913
DB 2022 AAGAGATTCCACCAAGTTGAAGAGCCT 2048
```

RESULT 6
US-09-881-578-3
Sequence 3, Application US/09881578
Publication No. US2003003966A1

GENERAL INFORMATION:
APPLICANT: Heiring, Thomas
APPLICANT: Johnstone, Brian
TITLE OF INVENTION: PROBES FOR CHONDROGENESIS
FILE REFERENCE: 27708/04013
CURRENT APPLICATION NUMBER: US/09/881,578
CURRENT FILING DATE: 2001-06-04
NUMBER OF SEQ ID NOS: 4
SOFTWARE: PatentIn version 3.0
SEQ ID NO 3

US-09-881-578-3
TYPE: DNA
ORGANISM: Homo sapiens

Query Match 34.9%; Score 703; DB 3; Length 2143;
Best Local Similarity 69.7%; Pred. No. 7.7e-206; Indels 27; Gaps 3;
Matches 1006; Conservative 0; Mismatches 410;

```
QY 172 TTCAGTTCTTAAACAATGCGCCAGGGTTTGTGACGTTCCGACGTAGCCATGACTTT 231
DB 10 TGCATTTCCAGAACCATGATGATGGTTTGTGACATTCAGGGATGTGGCCTGACCTTC 69
QY 232 TCTCAGAGAGTGGGCTGTCTGAACTTCTCTCAGAGGACCTGTACTGGGACGTGATG 291
DB 70 TCTCAGAGAGTGGGAAATGCTGACCTGCTCAGAGGACCTGTATCGTGATGTATG 129
QY 292 CTGGAGAACTACAGTAACCTTGCTCACTGATTTGAGTC--AGCATATGAAATAAG 348
DB 130 TTGGAGAACTATATGTAATCTTGTTGCTCAGATTTGAGTCAGAAACGTAATGACCAA 189
QY 349 AGTTTACCTACAGAAAAAACATTCATGAAATAAGGGCTTCAGAAAGAAATTCAGATA 408
DB 190 AAAATATTTTGAAGAAATGATATTTTGAATATATTTTCCAGTGGGAGATGAAGGAC 249
QY 409 AAGAGTAATCCCTTGGCC-----GTACTGATATGTGAAGT 447
DB 250 AAAAGTAAACCCCTTGGCCCTTGAGGATCAATCTTCAGAAATATTTGGAATGCAAAAGC 309
QY 448 AGCCTTGAAGAACACAGGCTCCAGAGGAGTATGTAATGTAATGATGATCAATTAAT 507
DB 310 ATATTCAGAGGACTTAAAGACATCAAGAGGATCTTCAATGTAATGTAATGATGACTAT 369
QY 508 GTCAAAAGGCTGTACTAGAGAAAGCACCCCTCCTAAGAACATCAGAGA--CATCAT 564
DB 370 GAAAAAATCCTTCTTACAGAAAAAGTAATCTTACTCAGATCAAGAAATTCATAAT 429
QY 565 AAGAGAAATTCCTTGAATGTAAAGACTGTGGGAAGGCTTTAGTGTGCTATCACTT 624
DB 430 ACAGAGAAATCCTATGTTGTAAAGAAATGTGGGAAGGCTTGCAATGATGCTCAAAACTT 489
QY 625 AGTCAACATCAAGAAATTCATCTGCTGAGAAACCTTAATGAAATGTAAGTAAGAG 684
DB 490 GTTCAACATGAGAACTCATACGCTGAAGAACATTTGAATGTAAGAAATGTGGGAAG 549
QY 685 GCCTTCCTTGGGGCAATCAGCTTCTCAACATCAAAAAATTCATCTGGGGAAGCC 744
DB 550 AATTAATTAAGTGGCTATCACTCAATGTGCAATGAGATTTCAATCTGTGTGAAACCC 609
QY 745 TACGAATGTAAGACTGTGGGAAGGCTTTTGCATGGGGCTCAAGCTCGTTATCTAAG 804
DB 610 TATGAGTGTAAAGAAATGTGGGAAGCCTTTAGCTGGGGATCAAGCCTTGTATTAACATGAG 669
QY 805 AAGATTCATATGAGTGAAGAAACCTTAATGTAATGTAAGAACTGTGGGAAGGCTTTGGCGCT 864
DB 670 AAGATTCATATGAGTGAAGAAACCTTAATGTAATGTAAGAAATGTGGGAAGGCTTTAGCTGT 729
QY 865 GGTGATGAGCTCACTCAGACCAAGATTCACACTGGGGAAGAACTACGAATGCATA 924
DB 730 GGTATCACTTACCAACATCAGAAATTCATATTTGTGTGAATCTTATTAATGAAG 789
QY 925 GACTGTGGGAAGCCTTTAGCCGTGTATTAATCTTATTCAGCAAGAGAAATTCATAGT 984
```


Db 790 GATGTGGAGGCGCTTTTGTGGGGCTCAAGCCTTGCTAAACATGATTAATTCATACA 849
Qy 985 GGGGAGAGCCTTACGAGTGTAAAGACTGTGGAGGCTTTTATTTGTGTTCAAGCTTC 1044
Db 850 GGTGAGAACTTATTAATGTAAAGATGTGGAGGCTTCAAGTCGTGCTATCAACTT 909
Qy 1045 ATTTCAGATAAAAGATTCAACAGGTGAGAAACCTATGAAATGTCAAGAAATGTGGAG 1104
Db 910 ACTCAGCATCAAGAAATCCATCTGTAAAGAACTTATGAAATGTAAATGTATGTGAAAG 969
Qy 1105 GCCTTACTCGAGTCAATTACCTTACCTCAGCATCAGAAATCCACACCGGTAGAGACCT 1164
Db 970 GCTTTTGTGTGGGCTATCACTTACTCGACATTCAGATTTTCATCTGTGTAGAAACCC 1029
Qy 1165 CACGAATGTAAAGAGTGTGGAGGCTTTCCGTGGGTTCAAGCTCTTTAAGCAGAG 1224
Db 1030 TATGAATGCAGAAATGTGGAGGCTTTTAAATTCGAGATCAAGTCTTATTCACATGAA 1089
Qy 1225 AGGATACATACGGGCGAGAGGCGCTTACAGTCAAGTCAAGATGTGGAGGCTTCAATTGT 1284
Db 1090 AATAATTCATACGTGTAGAGAACTTATGAAATGTAAAGATGTGAAAGGCTTTTGTGTCG 1149
Qy 1285 GGCATACCTCCTCAGCAGCAGAGATCCACACAGCGGAAACCCCGTATTAATGTAAAG 1344
Db 1150 GGCATACCTTCTCAACATCAGAAATCCATACCTGTGAGAAACCTTTTGAATGTAAAG 1209
Qy 1345 GAGTGTGGAGGCTTTCATTTATGTAGTCAAGCTTCGTGAAACATGAGAAATTCATACC 1404
Db 1210 GAATGTGGAGGCTTTTGTGGGTTCAAGCCTTGTAAACATGAGAGATGCTATACCT 1269
Qy 1405 GGGGTGAAACCTTATGTGGGTGTACAGAAATGTGGAGAGCTTTTATGTACAGGCTTCACT 1464
Db 1270 GGTGAGAAATCCATTAATGTAAAGATCGGAGAAACCTTTTGTATGTGTATCAACTT 1329
Qy 1465 ACACAACTACAGAAACGACAGTGGGCGGAAATCCTTACGAATGTAAAGAGTGGGAGAG 1524
Db 1330 ACTCGACATCAGGTATTTTCACTGTGTAGAAACCTTATGAAATGTAAAGATGTGGAGAG 1389
Qy 1525 GCATGTACACCTTAACCATCTCCGAGAACATCAGAGATCCACAAAGTTGAAGAGCC 1584
Db 1390 GCTTTTAAATTTGTAGTCAAGCCTTGTTCACATGAAAGATCCATCAGGAGGAGAAACC 1449
Qy 1585 TTT 1587
Db 1450 TAT 1452

RESULT 7
US-10-623-914-3
Sequence 3, Application US/10623914
Publication No. US20050075495A1
GENERAL INFORMATION:
APPLICANT: Hering, Thomas M.
TITLE OF INVENTION: PROBES FOR CHONDROGENESIS
FILE REFERENCE: 27708/03905
CURRENT APPLICATION NUMBER: US/10/623, 914
CURRENT FILING DATE: 2003-07-21
NUMBER OF SEQ ID NOS: 13
SOFTWARE: PatentIn version 3.1
SEQ ID NO 3
LENGTH: 2143
TYPE: DNA
ORGANISM: Homo sapiens
US-10-623-914-3

Query Match 34.9%; Score 703; DB 9; Length 2143;
Best Local Similarity 69.7%; Pred. No. 7.7e-206;
Matches 1006; Conservative 0; Mismatches 410; Indels 27; Gaps 3;

Qy 172 TTCAGTTCTTAAACAAATGGCCGAGGGTTTGTGACCTTGGCCGACGTACCATGACTTT 231
Db 10 TCGAATTCAGAAACATGACTGATGGGTTGTGATCACTTCAAGGAGATGTGGCATCGACTTC 69

Qy 232 TCTCAGAGAGTGGGCTGTCTGAACCTCTCTCAAGAGGACCTGTACTGGGACGTGATG 291
Db 70 TCTCAGAGAGAGTGGAAATGCTTGAACCTCTCTCAAGAGGACCTGTACTGGATGTATG 129
Qy 292 CTGAGAACTACAGTAACTTGGTCTCACTGATTTGAGTCC--AGCATATGAATAATAG 348
Db 130 TTGAGAACTATAGTAACTTGGTCTCACTGATTTGAGTCCAAATGATGAGACCAA 189
Qy 349 AGTTTACCTACAGAAATAAACATTCATGAATTAAGGCTTCCAAAGAAATTCAGATGA 408
Db 190 AAAATATTTTCAAGAAATGATATTTTGAATAATTTTCCAGAGTGGAGATGAAGAC 249
Qy 409 AGAATTAATCCCTTGGCC-----GTAATGATATGTAAGT 447
Db 250 AAAATTAACCTTGGCTTGGAGGATTCATCTTCAAGAAATTTGAAGTCAAAAGC 309
Qy 448 ACGTTGAAAGACACACAGGCTCCAGAGGAGTATGTCATACAGATGATCATTAAT 507
Db 310 ATATTGAGGAGCTTAAAGAGACATCAGAGGATCTTCACTCAATGATATATAGCTAT 369
Qy 508 GTCAAAAGGCTGTCTATAGAGAGGCAACCTCTCTAGAACATCAGAGA--CATCAT 564
Db 370 GAAAAATACCTTCTTACAGAAAGATTAATCTTCTTACCATCAATCAAGAAATCATAT 429
Qy 565 AAGAGAAATCTTGAATGTAAAGACTGTGGAGAGGCTTTATGTGTGCTATCACTT 624
Db 430 ACAGAGAAATCCTATGTTTGTAAAGAAATGTGGAGGCTTGCAGCATGCTCAAAACTT 489
Qy 625 AGTCAACATCAGAAATCCATCTGTGAGAAACCTTATGAAATGTAAAGATGTAAAG 684
Db 490 GTTCAACATGAGAACTATACACTGAAAGACCTTGAATGTAAAGATGTGGAGAG 549
Qy 685 GCTTCCGTTGGGCAATCAGCTTACATCAATCAAAATTCATATCTGGGAGAGCC 744
Db 550 AATTATTTAAGTCTATCACTCAATGTGATCAGAGATTTTCAATCTGTGTAGAAACC 609
Qy 745 TACGATGTAAAGACTGTGGAGAGGCTTTGATGGGCTCAAGCTCTGTTATTCATAG 804
Db 610 TATGAGTGTAAAGAAATGTGGAGAACTTTAGCTGGGATCAAGCTTGTATTAACATAG 669
Qy 805 AGGATTCATCTGTGAAAAACCTATGAAATGTAAAGACTGTGAAAGGCTTTGGGCGT 864
Db 670 ABAATTCACACTGTGAGAAACCTTATGAAATGTAAAGATGTGGAGAGGCTTTAGTGT 729
Qy 865 GGTGATGAGCTCACTCAGACCAAGATTCACACTGGGAGAAAGACTACGATGCAAA 924
Db 730 GGCATCACTTACCAACATCAGAAATTCATATTTGGTGTGAATCTTATTAATGTAAAG 789
Qy 925 GACTGTGGAGAGACTTTAGCCGTGTATTAATCTTATTCAGACAAAGAAATTCATAGT 984
Db 790 GAATGTGGAGGCTTTTGTGGGCTCAAGCCTTGTAAACATGAAATTAATTCATCA 849
Qy 985 GGGGAGAGGCTTACAGAGTAAAGACTGTGGAGAGGCTTTTATTTGTGTTCAAGCTTC 1044
Db 850 GGTGAGAACTTATTAATGTAAAGATGTGGAGAGGCTTCAAGTCGTGCTATCAACTT 909
Qy 1045 ATTTCAGATAAAAGATTCAACAGGTGAGAAACCTTATGAAATGTCAAGAAATGTGGAG 1104
Db 910 ACTCAGCATCAAGAAATCCATCTGTAAAGAACTTATGAAATGTAAATGTATGTGAAAG 969
Qy 1105 GCCTTACTCGAGTCAATTACTTACTCAGCATCAGAAATCCACACCGGTGAGAGGCTT 1164
Db 970 GCTTTTGTGTGGGCTATCACTTACTCGACATTCAGATTTTCATATCTGTGTAGAAACCC 1029
Qy 1165 CACGAATGTAAAGAGTGTGGAGGCTTTCCGTGGGTTCAAGCTCTTTAAGCAGAG 1224
Db 1030 TATGAATGCAGAAATGTGGAGGCTTTTAAATTCGAGATCAAGTCTTATTCACATGAA 1089
Qy 1225 AGGATACATACGGGCGAGAGGCGCTTACAGTCAAGTCAAGATGTGGAGAGGCTTCAATTGT 1284
Db 1090 AATAATTCATACGTGTAGAGAACTTATGAAATGTAAAGATGTGAAAGGCTTTTGTGTCG 1149

QY 1285 GGCCTATCACCTCAGCAGAGAGATTCACACAGCGGAAACCCCGTAATTAATGTAAG 1344
 |||||
 DB 1150 GGCCTATCACCTTCTCAGACATCAGAAATCCATATCTGTGAGAAACCTTTGTAATGTAAG 1209
 |||||
 QY 1345 GAGTGTGGGAAGGCTTTCATTTATGATGAGGCTGTGAAACATAGAGAAATTCATACC 1404
 |||||
 DB 1210 GAATGTGGGAAGGCTTGTAGTGGGTTCAGGCTTGTAAACATAGAGAGATTCATACC 1269
 |||||
 QY 1405 GGGGTAAACCCCTATGGGTGTACAGATGTGGGAAGAGCTTTAGTACAGGCTCAGCTT 1464
 |||||
 DB 1270 GGTGAGAAATCCCATGATATGTAAGATGCGGAAAGACCTTTGTAGTGGGTATCAGCTT 1329
 |||||
 QY 1465 ACACAACATCAGAAACGACAGTGGGCGAAATCCTACGATGTAAAGAGTGGGAAAG 1524
 |||||
 DB 1330 ACTGCATCATGAGTATTTTACACATGCTGATGAGAAACCTATGATATGTAAGAAATGTGGGAAAG 1389
 |||||
 QY 1525 GCATGTATACCATTAACCATCTCCAGAGACATCAGAGATTCACAAACAGTTGAAGAGC 1584
 |||||
 DB 1390 GCTTTTAATTTGGATTCAGGCTTGTTCACATGAAAGAAATCCATACAGGGGAGAAACCC 1449
 |||||
 QY 1585 TTT 1587
 |||||
 DB 1450 TAT 1452

RESULT 8

US-10-094-749-1120

; Sequence 1120, Application US/10094749

; Publication No. US20030219741A1

; GENERAL INFORMATION:

; APPLICANT: ISOGAI, TAKAO
 ; APPLICANT: SUGIYAMA, TOMOYASU
 ; APPLICANT: OTSUKI, TETSUJI
 ; APPLICANT: WAKAMATSU, AI
 ; APPLICANT: SATO, HIROYUKI
 ; APPLICANT: ISHII, SHIZUKO
 ; APPLICANT: YAMAMOTO, JUN-ICHI
 ; APPLICANT: ISONO, YUUKO
 ; APPLICANT: HIO, YURI
 ; APPLICANT: OTSUKA, KAORU
 ; APPLICANT: NAGAI, KEIICHI
 ; APPLICANT: IRIE, RYOTARO
 ; APPLICANT: TAMECHIKA, ICHIRO
 ; APPLICANT: SEKI, NAOHICO
 ; APPLICANT: YOSHIKAWA, TSUTOMU
 ; APPLICANT: OTSUKA, MOTOKUKI
 ; APPLICANT: NAGAHARI, KENJI
 ; APPLICANT: MASUHO, YASUHIKO
 ; TITLE OF INVENTION: NOVEL FULL-LENGTH cDNA
 ; FILE REFERENCE: 084335/0160
 ; CURRENT APPLICATION NUMBER: US/10/094,749
 ; PRIOR FILING DATE: 2002-03-12
 ; PRIOR APPLICATION NUMBER: 60/350,435
 ; PRIOR FILING DATE: 2002-01-24
 ; PRIOR APPLICATION NUMBER: JP 2001-328381
 ; PRIOR FILING DATE: 2001-09-14
 ; NUMBER OF SEQ ID NOS: 3381
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 1120
 ; LENGTH: 1945
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; US-10-094-749-1120

Query Match 30.7%; Score 618; DB 6; Length 1945;

Best Local Similarity 65.8%; Pred. No. 1,56-179;

Matches 956; Conservative 0; Mismatches 470; Indels 26; Gaps 3;

QY 145 GGAACCCCGAGAAGACTGATCAGTTCTT--CAGTTTAAACCATGAGCCAGGGTTGG 202
 |||||
 DB 217 GGAACCCAGAAAGAGACTGATCATTCTTGAGAGCTTAAGAAACATGGCTCGGAAATTAAG 276
 |||||
 QY 203 TGACGTTGCCGACGTAGCCATAGACTTTTCTCAGAGAGAGTGGGCTGTCTGAACCTCTG 262

DB 277 TGAATGTACAGGATATGTGCAATTTGCTCTCCAGAGAGAGGGAATGCTGGACTCTG 336
 |||||
 QY 263 CTCAGAGGAGCTGTACTGTGGACGTGATGCTGGAAGATCACTAGTAATCTGTCTCACTGG 322
 |||||
 DB 337 CTCAGAGAGATTTGTAAGAGATGTGATGTTGAGAGACATATAGCAACTTGTATCACTAG 396
 |||||
 QY 332 ATTGAGAGTCAAGATATGAAATTAAGAGTTTACCTACAGAAAAAAACATTCATGAATTA 382
 |||||
 DB 397 ACTTGCTTCAAGAGTGTGCAAGTAAGAGCTTATCTCCAGAAAAAGAACTTATGAAGACAG 456
 |||||
 QY 383 GGGCTTCCAAAAGAAATTCAGATAGAGAGATTAATCCCTTG----- 424
 |||||
 DB 457 AATTATCCCAATGGGAATAGAGTACAGACTTGAATCTTGTAAGAGTCCATT 516
 |||||
 QY 425 ---GCCGTACTGTGATATGTGAAGTACGCTTGAAGAACACAGAGCTCCAGAGGAGGT 481
 |||||
 DB 517 CCAAGGATTTTGTGAAGCCAAAGCCAGATGAGAGAGACAAAGAAATCAGAAAGAT 576
 |||||
 QY 482 ATGTCAATCAGATGATCATCAATTAATGTCAAAAGGCTGTCTATAGAGAGGACCCCTC 541
 |||||
 DB 577 ATTTCAAGCAGAGAGATGATCATATATGACAAATGTCCATTTTCAACGACATCTTACT 636
 |||||
 QY 542 CTAGAACATATCAGAG---CATCATAGAGAGAAATTCCTTGAATGTAGAGACTGTGGAG 598
 |||||
 DB 637 TATCTCAACATTCAGAGATGCTATCTCAGAGAAACCTATTAATGTAGAAATGTGGAG 696
 |||||
 QY 599 AGGCTTTAGTGTGCTATTCACATTAAGTCAACATCAGAAAAATCCATCTGTGTAGAAAC 658
 |||||
 DB 697 AAGCTTCAGAGAGAGCTCAACCTAACCAACATCAAGTATCTATCTGTGAAAGAAAC 756
 |||||
 QY 659 CTTATGATGTAAAGATGTAAAGAGGCTTCGTTGGGCAATCAGCTTACTCAATC 718
 |||||
 DB 757 CTTATGATGTAAAGCAATCGGGAAGGCTTTATGTGTATTCACAACTCAGCTTCAATC 816
 |||||
 QY 719 AAAAAATCTACTGTGGGGAAGGCTTCAGATGTAAAGCTGTGGAGGCTTTCAGAT 778
 |||||
 DB 817 AGAGACTTCATATCTGTGAGAAACCTATGCAATGTAAAGATGTGGAGGCTTATCTC 876
 |||||
 QY 779 GGGGCTCAAGCTCGTTATTCATTAAGAGAGATTCATCTGTGTAAAGAAACCTATGAATGTA 838
 |||||
 DB 877 AAGCTCAACATTTATTTATCATCATAGAAATTCATCTGTGAAAGAAACCATTAATATGTG 936
 |||||
 QY 839 AAGACTGTGAAAGGCTTTCGCGGTGTGATGAGCTCACTACGACCAAGATTCACA 898
 |||||
 DB 937 AAGATGTGGAAGAGCTTTATTCGATCTCAACAATTCACGACATCAAAAAAGTTCATA 996
 |||||
 QY 899 CTGGGAGAAAGACTACGATGCAAGAGCTGTGGAGAGCTTACCGGTGTATTAAC 958
 |||||
 DB 997 CTGGTGAAGAACTTATGTAAATGTAAAGATGTGGAGGCTTTACTCAGAAATTCACAC 1056
 |||||
 QY 959 TTATTCAGCAACAAGAAATTCATAGTGGGAGAGGCTTACGATGTAAAGACTGTGGGA 1018
 |||||
 DB 1057 TTACACTACACAGAGACTTCATCTGTGGAAGAGCTCATAGATGTAAAGATGTAGGA 1116
 |||||
 QY 1019 AGGCTTTATTTGTGTGTTCAAGCTTCATTCAGCAATTAAGAAATTCACACAGGTGAGAAC 1078
 |||||
 DB 1117 AGGTCTTTACTAGCTCTCAACAATTAATCTGCAATTAAGAAATTCATTCACGAGTAAGAAC 1176
 |||||
 QY 1079 CCTATGAATGTCAAGAAATGTGGAGGCTTTTACTGAGTCAATTCCTATCTACAGATC 1138
 |||||
 DB 1177 CCTATGAATGTAAAGAAATGTGAAAGGCTTTTATTTGTGAGCTCACAGCTTCTCAACATC 1236
 |||||
 QY 1139 AGAAGATTCACACCGGTGAGAGGCTTCAGAAATGTAAAGATGTGGAGAGGCTTTCGCT 1198
 |||||
 DB 1237 AGAAATTTCTATATGGGAGAAACCATATGATGTAAAGAAATGTGAGAGGCTTTATTC 1296
 |||||
 QY 1199 GGGGTTCAAGCTCGTTAAGCAGAGAGATCATACGGGCGAGAAACCGTACAGTGA 1258
 |||||
 DB 1297 GGGGCTCACTAGTGAACAACATCAGAGATTCATCTGTGTGAAGAAACCTTATTAATGTG 1356
 |||||
 QY 1259 CAGAAATGTGGAAGGCTTCAATTTGTGCTATCACTCACTCAGCAAGAGAAATCCACA 1318
 |||||

Db 1357 AAGAAATGAGGAAGGCTTATCCGAGGCTCAAACTTATCAACACGAGAAATTCACA 1416
Qy 1319 CAGCGGAAACCCGTAATAATGTAAGAGTGGGAGGCTTCAATTATGATGAGCC 1378
Db 1417 CCAATGAAAGCCCTATGATGTAAGATGTGGAAGATGTTTATGTCATGCTCAAC 1476
Qy 1379 TCGTAAACATGAGGAATTCATCCGAGGAGAAACCCATGAGGATGACGAATGAGGA 1438
Db 1477 TTACTCAACATGAGGAATTCACACTGGGAGAAACCTTATCAATGTAAAGAAATGGGA 1536
Qy 1439 AGAGCTTATGATCAAGGCTTATCAACATCAAGAAACGACAGTGGGAGAAAT 1498
Db 1537 AAGCTTTATATGATGATCACTCTTACACGACACAGAGATTCATCTGGTGAAGAAC 1596
Qy 1499 CTTAGATGTAAGAGTGGGAGGAGCATGTAAACCACTTAACATCTCCAGAAATC 1558
Db 1597 CCTATGAATGTAAAGATGTGGAAGAAACCTTATGCTGTGCTCAGAACTTACTCAACATG 1656
Qy 1559 AGAGATCCACA 1570
Db 1657 AGCGAATTCACA 1668

RESULT 9

US-10-108-260A-2359
; Sequence 2359, Application US/10108260A
; Publication No. US20040005560A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: No. US20040005560A1 full length cDNA
; FILE REFERENCE: H1-A0106
; CURRENT APPLICATION NUMBER: US/10/108,260A
; CURRENT FILING DATE: 2002-03-27
; NUMBER OF SEQ ID NOS: 5458
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2359
; LENGTH: 2235
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-108-260A-2359

Query Match 30.5%; Score 615.4; DB 6; Length 2235;
Best Local Similarity 75.3%; Pred. No. 1e-178;
Matches 766; Conservative 0; Mismatches 251; Indels 0; Gaps 0;

Qy 559 CATCATAGAGAGAAATTCCTTGAATGTAAAGACTGTGGAGGCTTTATGCTGGCTAT 618
Db 894 CATACTGGTGAAGAAACCTATGATGTAAAGATGTGGAGAACCTTTAGCTGGGATCA 953
Qy 619 CAACCTTATGATCAATCAAGAAATCCATATCTGTGAGAAACCTTATGAATGTAAAGATGT 678
Db 954 AGCTTTGTTAAACATGAGGAATTCACACTGTGTGAGAAACCTTATGAATGTAAAGATGT 1013
Qy 679 AAGAGGAGCTTCGTTGGGGAATCAGCTTACTCAACATCAAAATTCATCTGGGAG 738
Db 1014 GGGAGAGGCTTTATGCTGTGCTATCACTTACCAACATCAAAATTCATCTGGTGTG 1073
Qy 739 AAGCCCTTACGAATGTAAAGACTGTGGAGGCTTTTCATGTGGGCTCAAGCCCTGTTAT 798
Db 1074 AATCTTAAATGTAAAGATGTGGAGGCTTTTGTGGGGCTCAAGCCCTGCTAAA 1133
Qy 799 CATTAAGAGATTCATATCTGTGAGAAACCTTATGAATGTAAAGACTGTGAGAAAGCTTT 858
Db 1134 CATGAGATTAATTCATACAGGTAGAAACCTTATTAATGTAAAGATGTGAGAGGCTTTC 1193
Qy 859 CGGCTGTGATGAGTCACTCAGACACAGAGATTCACACTGGGAGGAAGACTATCGAA 918
Db 1199 AATCGTGGCTATCACTTATCTCAGACTCAAGAAATCCATCTGTGTAAGAAACCTTATGAA 1253
Qy 919 TGCATAAGACTGTGGAGAACTTTAGCCGTGTGTATTAACCTTATTCAGACAAGAAAT 978
Db 1254 TGTAAATATGTGGAAGGCTTTTGTGGGCTATCACTTACTGACATCAGATATTT 1313

Qy 979 CATAGTGGGAGGAAGGCTTATCGAGTGTAAAGACTGTGGAGGCTTTTATTGTGTCTCA 1038
Db 1314 CATACTGGTGAAGAAACCTTATGATGTCAAGAAATGTGGAGGCTTTTATTTGGGATCA 1373
Qy 1039 AGCTCATATTCAGATTAAGAAATTCACACAGTGAAGAAACCTTATGAATGTCAAGATGT 1098
Db 1374 AGCTTATTAACAATGAAGAAATTCATCTGTGTGAAGAAACCTTATGAATGTAAAGATGT 1433
Qy 1099 GGGAGGCTTTTATCTCGAGTCAATTTACTTACGATTCAGAAATTCACACCGGTGAG 1158
Db 1434 GGAAGGCTTTTATGCTGTGATCACTTCTCAACATCAAGAAATTCATATCTGTGAG 1493
Qy 1159 AAGCTTACGAATGTAAAGAGTGTGGAAGGCTTTCTGAGGCTTCAGAGCTCGTTAAG 1218
Db 1494 AATCTTTGAATGTAAAGAAATGTGGAAGGCTTTTATGTTGGGTTCAAGCTTGTAAA 1553
Qy 1219 CACGAGAGATATCATACGGGAGAGGCGTAACAATGTCACAAATGTGGAGGCTTTC 1278
Db 1554 CATGAGAGATTCATATCTGTGAGAAATCCCATGAATGTAAAGATGTGGAAGGCTTTT 1613
Qy 1279 AATTTGGCTATCACTTACATCAGACGAGAAATCCACAGGCGAAACCCGTAATAA 1338
Db 1614 TGTAGTGGGTATCACTTACTGACATCAGATGTATTCACACTGTGAGAAACCTTATGAA 1673
Qy 1339 TGTAAAGAGTGTGGAGGCTTTTATGATGATGAGCTCTGTGAACATGAGAGAAAT 1398
Db 1674 TGTAAAGAAATGTGGAGGCTTTTATGATGATGATGAGCTTGTTCACATTAAGAAATC 1733
Qy 1399 CATACGGGAGTGAAGAAACCTATGAGTGTAAAGAAATGTGGAGAAAGCTTTATGATCAGGCAAT 1458
Db 1734 CATACAGGAGGAAGAAACCTTATGAATGTAAAGATGTGGAGGCTTTTATGATGATGAT 1793
Qy 1459 CAGCTTACAAATCATCAAGAAACGACAGTGGGCGAAATCTTACGAATGTAAAGATGTC 1518
Db 1794 CACCTTACTCAACATCAAGAAATTCATACCGGTGAGAAACCTTTCAATGTAAAGAAATGT 1853
Qy 1519 GGGAGGAGTGTAAACCATCTCGAAGAAATCATGAGAGATCCACAACAGT 1575
Db 1854 GGGAGGCTTCAAGTGGGCTTCAAGCTTATGATGATGAGAGATTCATATCTAAT 1910

RESULT 10

US-10-450-763-25299
; Sequence 25299, Application US/10450763
; Publication No. US20050196754A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 790CIP3/US
; CURRENT APPLICATION NUMBER: US/10/450,763
; CURRENT FILING DATE: 2003-06-11
; PRIOR APPLICATION NUMBER: PCT/US01/08631
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 09/649,167
; PRIOR FILING DATE: 2000-08-23
; NUMBER OF SEQ ID NOS: 60736
; SOFTWARE: Custom
; SEQ ID NO 25299
; LENGTH: 1900
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIMILAR
; LOCATION: (28)..(1464)
; OTHER INFORMATION: 60% homologous to Homo sapiens zinc finger protein, accession
; OTHER INFORMATION: number X78933, Smith-Waterman Score=1602.
US-10-450-763-25299

Query Match 30.0%; Score 603.6; DB 9; Length 1900;
Best Local Similarity 70.3%; Pred. No. 4.2e-175;
Matches 823; Conservative 0; Mismatches 344; Indels 3; Gaps 1;

```
QY 429 TAACTGATATGTGAAGTACGCTTGAAAGACACAGCGCTCCAGAGGAGTATGTCAA 488
DB 51 TAAATGGAGTGAAGAAAGCATATTCAGAGGACATAAAGAGCATCAAGAGGATACTTCAG 110
QY 489 TCAGATGATCATCAATTAATGTCAAAGCGCTGCTACTAGAGAGGCAACCCCTCTTGAAC 548
DB 111 TCAAAATGATATCAGCTATGAAAAAATACCTTCTACAGAAAAAGTAAATCTCTTCTCC 170
QY 549 ACATCAAGACA---CATCATAGAGGAATCCCTTGATGTAAAGACGCTGGGAGGCGCTT 605
DB 171 ACATCAAGAAATTCATTAATACAGAAATCTAATGTTTGTAAAGAAATGTGGAGGCGTTG 230
QY 606 TAGCTGCGCTATCACTTAATAGTCAATCAAGAAATCACTACTGCTGAGAAACCTTATGA 665
DB 231 CAGTCACTGCTCAAACTTGTTCATCATGAGAACATCAACGCTGAGAAAAACACTTGA 290
QY 666 ATGTAAAGATGTGAAGAGCGCTTCGTTGGGCAATCAGCTTACTCAATCAATCAAAAAT 725
DB 291 ATGTAAAGATGTGGAGAAATTAATTAAGTCCATCACTCAATGTGCATCAGAGATT 350
QY 726 TCATCTGGGAGAGAGCGCTTACGATGTAAAGACGCTGGGAGGCGCTTTCGATGGGCGTC 785
DB 351 TCATCTGGGAGAGAGCGCTTACGATGTAAAGACGCTGGGAGAGCGCTTTCGATGGGAGTC 410
QY 786 AAGCTCTGTTATTCATTAAGAGATTCATCTGCTGAGAAAAACCTTATGAATGTAAAGACTG 845
DB 411 AAGCTCTGTTAAACATGAGAGATTCACCTGCTGAGAAAAACCTTATGAATGTAAAGATG 470
QY 846 TGGAAAGCGCTTCGCGCTGTGTATGAGTCACTCAGCACAGAGATTCCACTGGGGA 905
DB 471 TGGGAGAGCGCTTATGCTGTGCTATCCTTACCACATCAAAAATTCATCTGCTGT 530
QY 906 GAAAGACTAGTGAAGAGCTGTGGAGAGCGCTTAAAGCGGTGATTAATCACTTATCA 965
DB 531 GAAATCTTAAATGTGAAGAAATGTGGAGAGCGCTTATTTTGGGCGCTCAAGCGCTTGTCA 590
QY 966 GCACAAGAGATTCATTAAGTGGGAGAGCGCTTACGAGTGAAGAGCTGTGGAGAGCGCTT 1025
DB 591 ACATGAGATTAATTCATCAGTGAAGAACTTAAATGTAAAGAAATGTGGAGAGCGCTT 650
QY 1026 TATTTGCTTCAAGCGCTTCAAGCATTAAGAAATTCACACAGTGTGAAGAAACCTTATGA 1085
DB 651 CAGTCTGTGCTATCAGCTTCAAGCATCAAGAAATTCATCTGTGAAGAAACCTTATGA 710
QY 1086 ATGTCAAGATGTGGAGAGCGCTTACTCGAGTCAATTAATTAATCACTCAGATCAGAGAT 1145
DB 711 ATGTAAATTAATGTGAAGAGCGCTTGTGTGGGCTATCACTTCACTCAGATCAGATATT 770
QY 1146 CCACACCGGTGAGAGAGCGCTCAGATGTGAAGAGTGTGGAGAGCGCTTTCGCTGGGCTTC 1205
DB 771 TCATTAATGTGAGAGAGCGCTTATGAATGCAAGAGATGTGGAGAGCGCTTTCATTTGCGGATC 830
QY 1206 GAGCGCTGTGAAGACAGAGAGATCACTACGCGGAGAGAGCGCTTCAAGTGTCAAGATG 1265
DB 831 AAGCTTATTAATCAATGAAGAAATTCATCTGTGAGAAAACTTAAATGTGAAGATG 890
QY 1266 TGGAGAGCGCTTCAATTAATGTGGCTATCACTCAGACAGAGAGATTCACACAGGCGA 1325
DB 891 TGGAGAGCGCTTATGATGTGCTATCACTTTCACATCAGAAAAATTCATCTGTGTA 950
QY 1326 AACCCGCTATTAATGTGAAGAGTGTGGAGAGCGCTTTCATTTATGATCGAGCGCTGTGA 1385
DB 951 GAAACCTTTTGAATGTGAAGAGTGTGGAGAGCGCTTTCATTTGAGGCTTCAAGCGCTGTGA 1410
QY 1386 ACATGAGAGATTCATTCGCGGCTGAAACCTTATGCGGTGTCAGATGTGTGAAGAGCTT 1445
DB 1011 ACATGAGAGATTCATTCGCGGCTGAAATCCATGAATGTGAAGAGCGGAAAGACCTT 1470
QY 1446 TAGTCAAGCGGCTCAGCTTCAAGATCAGAAAAAGCAGACAGTGGGCGGAAATCTTCA 1505
DB 1071 TTGTAGTGGTATCACTTACTCGACATCAGGTATTTTCACTGTGTGAAGAAACCTTATGA 1130
```

```
QY 1506 ATGTAAAGAGTGGGAGAGCATGTAAACCACTTAACCACTCCAGAGAAATCAGAGAT 1565
DB 1131 ATGTAAAGAAATGTGGAGAGCGCTTTCATTTATGTGTGATCAAGCGCTTGTTCACATGAAGAAAT 1190
QY 1566 CCACAACAGTTGAAGAGCGCTTTTGAACGCA 1595
DB 1191 CCATACAGGAGAGAAACCTTATGAAGATGA 1220

RESULT 11
US-09-764-864-567
; Sequence 367, Application US/09764864
; Patent No. US20020132753A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PRZ23
; CURRENT APPLICATION NUMBER: US/09/764,864
; PRIOR APPLICATION DATA: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1792
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO: 567
; LENGTH: 774
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (745)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-764-864-567

Query Match 29.5%; Score 593.6; DB 3; Length 774;
Best Local Similarity 91.4%; Pred. No. 3.1e-172;
Matches 690; Conservative 5; Mismatches 25; Indels 35; Gaps 5;

QY 1 CGCTTGTGACCGGTGACGCAACCGCTGTCTCCGCGCATGTCGCGGAGCGCAGCATCT 60
DB 18 CGCTTGTGACCGGTGACGCAACCGCTGTCTCCGCGCATGTCGCGGAGCGCAGCATCT 77
QY 61 TCAGAAAGAGCATCCCGGAGAGAGAGAGATTCGTTAAACATCTTAAGTCC----- 111
DB 78 TCAGAAAGAGCATCCCGGAGAGAGAGAGAGATTCGTTAAACATCTTAAGTCC 137
QY 112 -----AGCTTACGCTCTCGGAATTTGTCTTCTTCAAGTGAAG 150
DB 138 AGTGAAGAAACGACACTGTAGCTCTAGCTCTCGGAATTTGTCTTCAAGTGAAG 197
QY 151 CCCGAGAGACTGATCAATTTCTTCAATTTTAAACAATGCGCGGATTTGTGACGTT 210
DB 198 CCCGAGAGACTGATCAATTTCTTCAATTTTAAACAATGCGCGGATTTGTGACGTT 257
QY 211 GCCGAGTGGCCATAGACTTTTCTCAGAGAGAGTGGGCGGTGTAACCTGTGCGAGG 270
DB 258 GCCGAGTGGCCATAGACTTTTCTCAGAGAGAGTGGGCGGTGTAACCTGTGCGAGG 317
QY 271 GACCTTACTGGGACCTGTAGTGTGAGAACTACATGTAATCTTGTCTCACTGATTTGAG 330
DB 318 GACCTTACTGGGACCTGTAGTGTGAGAACTACATGTAATCTTGTCTCACTGATTTGAG 377
QY 331 TCAGCATATGAATAATGAAGTTTACTTACAGAAAAAAACATTCATGAATTAAGGCTTCC 390
DB 378 TCAGCATATGAATAATGAAGTTTACTTACAGAAAAAAACATTCATGAATTAAGGCTTCC 437
QY 391 AAAAGGAATTCAGATGAAGAGTAAATCCCTTGGCGGTAACTGATGTGAAGTGAAG 450
DB 438 AAAAGGAATTCAGATGAAGAGTAAATCCCTTGGCGGTAACTGATGTGAAGTGAAG 497
QY 451 CTTGAAGAGCAAGAGGCTTCAGAGGAGTATGTCAATCAGATGATCAATTAATGTC 510
DB 498 CTTGAAGAGCAAGAGGCTTCAGAGGAGTATGTCAATCAGATGATCAATTAATGTC 557
QY 511 AAAAGCGCTGCTTATGAAGAGGACCCCTCTTGAACAATCAGAGACATCATTAAGAG 570
```



```

Db 1121 CATCAGAAATTCATACCGGTGAGAAACCATATATATGTAATGAAGTGGTAAGGCTTT 1180
Qy 1027 ATTGTGTTCAAGCTTCACTTACGATTAAGAATTCACACAGGTGAGAAACCTATGAA 1086
Db 1181 AATTTTGGATCGACCTTACTGACATCAAGAATTCATACGTGAGAAACCTTATGTA 1240
Qy 1087 TGTCAAGATGTGGAGAGGCTTTACTCGAGTCAATTAACCTTACTAGCATGAGAGATC 1146
Db 1241 TGTAAAGATGTGGAGAGGCTTTTATATGTGCTGAGATCTCACTAGCATGAGAGAT 1300
Qy 1147 CACACCGGTGAGAGGCTTCAAGATGTAAGAGTGTGGAGAGGCTTTGCTGGGTTGC 1206
Db 1301 CACACTGGTGAAGAAACCTATAGTGTAAAGAGTGTGAAGAACCTTTAGAAAGTGTCA 1360
Qy 1207 ACCCTGTAAACGACAGAGATACATAGGGGAGAAACCGTACAAAGTGCACAGATGT 1266
Db 1361 AACTTATTCACATCAAAAGAAATGCACTGAGAGAAACCTTATGATGTAAGGATGT 1420
Qy 1267 GGAAGGCTTCAATTGTGGCTATCACTCACTCAGACAGAGAAATCCAACAGGCGAA 1326
Db 1421 GGAAGACCTTTATAGTGTGTTCAGACCTTACTCAACATCAAGAAATTCATCTGTGAG 1480
Qy 1327 ACCCGTATTAATGTAAGAGTGTGGAGAGGCTTTCAATTATGATCGAGCTCGTAAA 1386
Db 1481 AAACCTATGAATGTAAAGAAATGTGGAGAGGCTTTGTAGTGTCAAAACCTTATCAA 1540
Qy 1387 CATGAGAAATTCATACCGGGGTGAACCTATGGGTGTAACGAATGTGGAGAGGCTTT 1446
Db 1541 CACCGCTATATCATATCGTGTGAAGACCTTATGAATGTAAGAAATGTGAAAGTCTTT 1600
Qy 1447 AGTCAGCGCCATCAGCTTACACAAATCAGAAACGACAGTGGGCGGAATCTTACGA 1506
Db 1601 AGTAGTGTTCAGCTTATATCGGCAACGAGAAATCACTGTGTAGAAACCTTATGAA 1660
Qy 1507 TGTAAAGTGTGGAGAGGATTAACCATTAACCATCTCGAGAAATCAAGAGATC 1566
Db 1661 TGTAAAGTGTGGAGAGGCTTTTATAGTGTCAAGGCTTACTGACATCAAGAAAT 1720
Qy 1567 CACAAAGTTGAAGAGGCTTTGAACGCACTAGC 1600
Db 1721 CA-TACAGGTGAGAAACCTTATGATGTAAGAAC 1753

```

RESULT 14

```

US-10-221-625-162
/ Sequence 162, Application US/10221625
/ Publication No. US2004003942A1
/ GENERAL INFORMATION:
/ APPLICANT: INCYTE GENOMICS, INC.
/ APPLICANT: HILLMAN, Jennifer L.
/ APPLICANT: BAUGHN, Mariah R.
/ APPLICANT: YUE, Henry
/ APPLICANT: LAL, Preeti
/ APPLICANT: LU, Dyrung Aina M.
/ APPLICANT: PATTERSON, Chandra
/ APPLICANT: AZIMZAI, Yalda
/ APPLICANT: BANDMAN, Olga
/ APPLICANT: TANG, Y. Tom
/ APPLICANT: WATHUR, Preeti
/ APPLICANT: SHAH, Purvi
/ APPLICANT: AU-YOUNG, Janice
/ APPLICANT: REDDY, Roopa
/ TITLE OF INVENTION: TRANSCRIPTION FACTORS
/ FILE REFERENCE: PP-0761 PCT
/ CURRENT APPLICATION NUMBER: US/10/221,625
/ NUMBER OF SEQ ID NOS: 214
/ SOFTWARE: PERL Program
/ SEQ ID NO 162
/ LENGTH: 1813
/ TYPE: DNA
/ ORGANISM: Homo sapiens

```

```

/ FEATURE:
/ NAME/KEY: misc_feature
/ OTHER INFORMATION: Incyte ID No. US2004003942A1 2154810CB1
US-10-221-625-162

Query Match 26.7%; Score 538.6; DB 7; Length 1813;
Best Local Similarity 69.0%; Pred. No. 5,4e-155;
Matches 793; Conservative 0; Mismatches 329; Indels 27; Gaps 3;

Qy 322 GATTTGAGTGCATATGAAATTAAGATTACTTACAGAAAGAAACATTCATGAATA 381
Db 377 GATTTGAGACCAAGTATGAGCAAGAAATTTATCTTTAGAAAGAAATGCAATTTATGAATA 436
Qy 382 AGGCGTTCCAAAAGAAATTCAGATGAAGAAAGTAAATCCCTTGGCCGTAAAC----- 432
Db 437 AATTTATCCAGTGAAGTAATGAAAGAAATTTGAAACCAATGGCTTTAAGGTCAT 496
Qy 433 -----TGAATATGTAAGTACGCTTGAAGAACCAACAGCGCTTCAGAGGAG 480
Db 497 TTAAGAAATGATTTGGAAATCCAACAGAAATTTGAAGACAGAGAGACCTCAAGAGGA 556
Qy 481 TATGTCAATCAGATGATCATCAATTATGTCAAAAGGCTGTACTAGAGAGGACCCCT 540
Db 557 TACTTCAAGTGTGAAAATGCAATCTGAAAAGTGTCTCTTACCAAGAAAGGCAATGT 616
Qy 541 CCTAGAACATCAGAGACATCA---TAAGAGAAATCCCTTGAATGTAAGACTGTGG 597
Db 617 GTTACTCCACATCAGAGACCTTCAATTTGTGTATTAACCTTATGAATGTAAGAAATGTGG 676
Qy 598 AAGCGCTTTATGTCGTGTCTATCACTTACTCAATCAAGAAATCCATCTGTGAGAA 657
Db 677 AAGGGGTTCAGAGTGCAGCAAGCTTACTTTTCAATCAAGAAATCATCTGTGAAAAA 736
Qy 658 CCTTATGAATGTAAGAAATGTAAGAGGCTTCCGTGGGGAATCAAGCTTACTCAACAT 717
Db 737 CCGTATGAATGTAAGAAATGTGGATGTGGCTTCAAGACACTGCACACTTACTGACAT 796
Qy 718 CAAAAATTCATATCGGGAGAGGCCCTTACGAATGTAAGACTGTGGAGAGGCTTTGCA 777
Db 797 CAGAGAAATTCATATCGTGTGAAGAACCTTATGAATGTAAGAAATGTGAGAAAGGCTTAACT 856
Qy 778 TGGGGCTCAAGCTCTGTTATTCATTAAGAGATTCATCTGTGTGAAAAAACCTTATGAATGT 837
Db 857 CGTGGCTATCACTTCTTCATCAAGAAATTCATCTGTGTGAGAAACCTTTTGATGT 916
Qy 838 AAGAAGTGTGAAGAGGCTTTGGGGGTGTGATGAGCTCACTCAGACAGAGATTCAC 897
Db 917 AAGGAATGTGAGAGGCTTTTATGTGGGTTCAGGCTTGTAAACATGAGAGATTCAT 976
Qy 898 ACTGGGAGAGAAAGCTTACGAATGCAAGACTGTGGAGAGCCTTTAGCCGTGTATAA 957
Db 977 ACTGGTGAAGAAATCCATGATGTAAGAAATGTGGAGAGACCTTTGTGTGGGTATCAA 1036
Qy 958 CTATTTCAGACAAAGAAATTCATTAAGTGGGAGAGGCTTTAGAGTGTAAAGCTGTGG 1017
Db 1037 CTTACTCGACATCAGTATTTTCACTGTGTGAAGAACCTTATGAAGTGAAGAAATGTGG 1096
Qy 1018 AAGGCTTTATTTGTGGTCAAGCCCTATTCAGCATTAAGAAATTCACAGGTGAGAA 1077
Db 1097 AAGGCTTTTAAATGTGTGATCAAGCCTTTGTTCACATGAAAGAAATTCATCAGGGAGAA 1156
Qy 1078 CCTATGAATGTCAAGAAATGTGGAGAGGCTTTTACTCGAGTCAATTAACCTTACTCAGAT 1137
Db 1157 CCTATGAATGTAAAGAAATGTGGAGAGGCTTTTATGTGTGATCACTTACTCAACAT 1216
Qy 1138 CAGAGATCAACCGGTGAGAGGCTTCAAGATGTAAGAGTGTGGAGAGGCTTTTGGC 1197
Db 1217 CAGAAATTCATATCCGTGTGAAGAACTTTCAATGTAAGAAATGTGGAGAGGCTTCAAGT 1276
Qy 1198 TGGGGTTGAGGCTGTGTAAGACAGAGAGATCATACGGGCGAGAGGCGTATCAAGTGC 1257
Db 1277 TGGGGTTCAAGCTATGTAAGACAGAGAGTCAATCTAATGAGAGTCTTATGATGT 1336

```


THIS PAGE BLANK (USPTO)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceeleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 6, 2006, 19:17:02 ; Search time 372 Seconds
(without alignments)
9628.459 Million cell updates/sec

Title: US-10-111-257-3

Perfect score: 2015
Sequence: 1 cgcctcgcaccggtgacgc.....gtgaagttcctaacttaa 2015

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1303057 seqs, 888780828 residues

Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

Issued Patents NA: *
1: /cgn2_6/ptodata/1/ina/1_COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/H_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
7: /cgn2_6/ptodata/1/ina/PP_COMB.seq:*
8: /cgn2_6/ptodata/1/ina/RE_COMB.seq:*
9: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1397.8	69.4	1429	3	US-09-023-655-689
2	703	34.9	2143	3	US-09-881-578A-3
3	520.6	25.8	2274	3	US-09-620-312D-154
4	510	25.3	2499	3	US-09-774-528-331
5	510	25.3	2499	3	US-10-120-988-331
6	502.8	25.0	3472	3	US-10-104-047-415
7	495.4	24.6	1694	3	US-10-104-047-1500
8	478.2	23.7	2925	3	US-09-620-312D-163
9	474.8	23.6	2241	3	US-10-104-047-693
10	461.6	22.9	1914	3	US-10-104-047-1464
11	460.6	22.9	3798	3	US-09-949-016-4204
12	460.6	22.9	156942	3	US-09-949-016-12227
13	460.6	22.9	156850	3	US-09-949-016-15946
14	456.6	22.7	2669	3	US-10-104-047-271
15	455.2	22.6	3839	3	US-09-949-016-485
16	451.2	22.4	2666	3	US-09-881-578A-1
17	446.2	22.1	2441	3	US-09-949-016-2756
18	446.2	22.1	19861	3	US-09-949-016-14498
19	445.4	22.1	2320	3	US-09-016-434-1054
20	445	22.1	2660	3	US-10-104-047-559
21	444.6	22.1	3078	3	US-10-104-047-622
22	437	21.7	2978	3	US-09-949-016-3823
23	436.6	21.7	45138	3	US-09-949-016-13027
24	433.6	21.5	3400	3	US-09-799-451-136

25	428.4	21.3	3138	3	US-10-104-047-83	Sequence 83, App1
26	422.8	21.0	2110	3	US-10-104-047-1778	Sequence 1778, App
27	422	20.9	15632	3	US-09-949-016-15565	Sequence 15565, A
28	418	20.7	2625	3	US-10-104-047-1026	Sequence 1026, App
29	416.6	20.7	2351	3	US-09-016-434-1337	Sequence 1337, App
30	416.6	20.7	2637	3	US-09-949-016-5623	Sequence 5623, App
31	416.6	20.7	27227	3	US-09-949-016-17365	Sequence 17365, A
32	415.2	20.6	2752	3	US-09-949-016-3563	Sequence 3563, App
33	415.2	20.6	18824	3	US-09-949-016-15305	Sequence 15305, A
34	412.6	20.5	3026	3	US-10-104-047-967	Sequence 967, App
35	410.4	20.4	1629	3	US-09-016-434-1336	Sequence 1336, App
36	407.2	20.2	1812	3	US-09-016-434-1407	Sequence 1407, App
37	403.4	20.0	128779	3	US-09-497-855A-38	Sequence 38, App1
38	403	20.0	3090	3	US-10-104-047-191	Sequence 191, App
39	402.6	20.0	2357	3	US-10-104-047-1701	Sequence 1701, App
40	400.2	19.9	2798	3	US-10-104-047-371	Sequence 371, App
41	400	19.9	3400	3	US-10-104-047-1037	Sequence 1037, App
42	397.6	19.7	2489	3	US-09-774-528-280	Sequence 280, App
43	397.6	19.7	2489	3	US-10-120-988-280	Sequence 280, App
44	397.6	19.7	2582	3	US-09-016-434-1399	Sequence 1399, App
45	396.6	19.7	2141	3	US-10-104-047-600	Sequence 600, App

ALIGNMENTS

RESULT 1
US-09-023-655-689
; Sequence 689, Application US/09023655
; Patent No. 6607879
; GENERAL INFORMATION:
; APPLICANT: Cocks, Benjamin G.
; APPLICANT: Susan G. Stuart
; APPLICANT: Jeffrey J. Selthamer
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL GENE
; TITLE OF INVENTION: EXPRESSION
; NUMBER OF SEQUENCES: 1508
; CORRESPONDENCE ADDRESS:
; ADDRESS: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/023.655
; FILING DATE: HERewith
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Zeller, Karen J.
; REGISTRATION NUMBER: 37,071
; REFERENCE/DOCKET NUMBER: PA-0001 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 855-0555
; TELEFAX: (650) 845-4166
; INFORMATION FOR SEQ ID NO: 689:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1429 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: TLYNMOTO2
; CLONE: 450088

US-09-023-655-689

Query Match 69.4%; Score 1397.8; DB 3; Length 1429;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 1411; Conservative 0; Mismatches 7; Indels 2; Gaps 1;

QY 595 GGGAGAGCCTTATGTCGTGGCTATCACTTATGTCACATCAGAAAAATCCATATGCTGTAG 654
Db 10 GGGAGAGCCTTATGTCGTGGCTATCACTTATGTCACATCAGAAAAATCCATATGCTGTAG 69
QY 655 AAACCTTATGATGTAAGAAATGTAAGAGGCTTCCGTGGGGCAATACCTTACTCA 714
Db 70 AAACCTTATGATGTAAGAAATGTAAGAGGCTTCCGTGGGGCAATACCTTACTCA 129
QY 715 CATCAAAAAATCATACTGGGGAGAAAGCCTTAAGATGTAAGAGCTGTGG--AAGGCTT 772
Db 130 CATCAAAAAATCATACTGGGGAGAAAGCCTTAAGATGTAAGAGCTGTGGGGAAAGGCTT 189
QY 773 TTGATGGGGCTCAAGCCTCGTTATTCATAGAGATTCATATCTGTGTAAGAAACCTATG 832
Db 190 TTGATGGGGCTCAAGCCTCGTTATTCATAGAGATTCATATCTGTGTAAGAAACCTATG 249
QY 833 AATGTAAGACTGTGGAAGAGCCTTTGGCGTGATGATGCTCACTCAGCAGAGAT 892
Db 250 AATGTAAGACTGTGGAAGAGCCTTTGGCGTGATGATGCTCACTCAGCAGAGAT 309
QY 893 TCCACACTGGGGAGAAAGCTAGAAAGAACTGTGGGAAAGCCTTTAGCCGTGT 952
Db 310 TCCACACTGGGGAGAAAGCTAGAAAGAACTGTGGGAAAGCCTTTAGCCGTGT 369
QY 953 ATAACTTATTCAGCAGAAAGAAATCATAGTGGGAGAAAGCCTTACGATGTAAGACT 1012
Db 370 ATAACTTATTCAGCAGAAAGAAATCATAGTGGGAGAAAGCCTTACGATGTAAGACT 429
QY 1013 GTGGGAAGCCTTATTTTGTGTTCAGAGCTCATTCAGCATATAAATAATTCACAGGTG 1072
Db 430 GTGGGAAGCCTTATTTTGTGTTCAGAGCTCATTCAGCATATAAATAATTCACAGGTG 489
QY 1073 AGAAACCTATGTAATGTCAAGAATGTGGGAAGCCTTTACTCGAGTCATTAACCTACTC 1132
Db 490 AGAAACCTATGTAATGTCAAGAATGTGGGAAGCCTTTACTCGAGTCATTAACCTACTC 549
QY 1133 AGCATCAGAAAGTCCACACCGGTGAGAGCCTTCAGAAATGTAAGAGTGTGGGAAGCCTT 1192
Db 550 AGCATCAGAAAGTCCACACCGGTGAGAGCCTTCAGAAATGTAAGAGTGTGGGAAGCCTT 609
QY 1193 TTGCGTGGGGTTCGAGCCTCGTTAAGCAGAGAGATCATATGGCGAGAAAGCCTTACA 1252
Db 610 TTGCGTGGGGTTCGAGCCTCGTTAAGCAGAGAGATCATATGGCGAGAAAGCCTTACA 669
QY 1253 AGTGCACAGAAATGTGGGAAGCCTTCAATTTGTGCTATCACTCAGCAGAGAGAA 1312
Db 670 AGTGCACAGAAATGTGGGAAGCCTTCAATTTGTGCTATCACTCAGCAGAGAGAA 729
QY 1313 TCCACACAGGCGAAACCCCGTATTAATGTAAGAGATGTGGGAAGCCTTCAATTTATGAT 1372
Db 730 TCCACACAGGCGAAACCCCGTATTAATGTAAGAGATGTGGGAAGCCTTCAATTTATGAT 789
QY 1373 CGAGCCTCGTGAAGATGAGAGATTCATTCGGGGTGAAGCCTTATGGGTATACGAAT 1432
Db 790 CGAGCCTCGTGAAGATGAGAGATTCATTCGGGGTGAAGCCTTATGGGTATACGAAT 849
QY 1433 GTGGGAAGAGCTTATGTCAGGCGCATCAGCTTACAAACATCAGAAAAAGCAGAGGGG 1492
Db 850 GTGGGAAGAGCTTATGTCAGGCGCATCAGCTTACAAACATCAGAAAAAGCAGAGGGG 909
QY 1493 CGAAATCTTACGAATGTAAAGAGTGGCGGAAGGCATGTAAACAATTAACATCTCCGAG 1552
Db 910 CGAAATCTTACGAATGTAAAGAGTGGCGGAAGGCATGTAAACAATTAACATCTCCGAG 969
QY 1553 AACATCAGAGGATCCACACAGATTTGAAGAGCCTTTTGAAGCAGATGAGCCGCTGTATCT 1612
Db 970 AACATCAGAGGATCCACACAGATTTGAAGAGCCTTTTGAAGCAGATGAGCCGCTGTATCT 1029

QY 1613 ATGTTTGCCTTCCACAGTTTGTATCCTGCAAGTCACTGCAAGTTCAAAAATATTTAAATG 1672
Db 1030 ATGTTTGCCTTCCACAGTTTGTATCTGACAGTCACTGCAAGTCAAAAATATTTAAATG 1089
QY 1673 GAAAATCCAGAAATTAAGAAATTTTAAATGTCAAATGTGTGCCCTTCTGAGTACCGTCA 1732
Db 1090 GAAAATCCAGAAATTAAGAAATTTTAAATGTCAAATGTGTGCCCTTCTGAGTACCGTCA 1149
QY 1733 TGAATTCCTGCTGTGCCGCTCCAGCGCGCGGGAGTGTAGTATCATCCTTGTGTCCAGC 1792
Db 1150 TGAATTCCTGCTGTGCCGCTCCAGCGCGCGGGAGTGTAGTATCATCCTTGTGTCCAGC 1209
QY 1793 ACATCAAGCTGTATACGCCAACCCACCTGCTAGTACTATGATAGCCGCTTTGTGTATCA 1852
Db 1210 ACATCAAGCTGTATACGCCAACCCACCTGCTAGTACTATGATAGCCGCTTTGTGTATCA 1269
QY 1853 GATCAACTATCCAGCATCAGATGCTGTGCCCAAGTATGCTTCACTTTGCTTAAAGT 1912
Db 1270 GATCAACTATCCAGCATCAGATGCTGTGCCCAAGTATGCTTCACTTTGCTTAAAGT 1329
QY 1913 GGCCCCAGAGCAGAGAGTGAATGATGCTGTGATTCGATATGCCAAGAGAGAGCCACA 1972
Db 1330 GGCCCCAGAGCAGAGAGTGAATGATGCTGTGATTCGATATGCCAAGAGAGAGCCACA 1389
QY 1973 AGTGCTTCCTTTTAATGAAAAAGTGAAGTTCTCACTT 2012
Db 1390 AGTGCTTCCTTTTAATGAAAAAGTGAAGTTCTCACTT 1429

RESULT 2
US-09-881-578A-3
; Sequence 3, Application US/09881578A
; Patent No. 6596855
; GENERAL INFORMATION:
; APPLICANT: Herling, Thomas M.
; APPLICANT: Johnstone, Brian
; TITLE OF INVENTION: PROBES FOR CHONDROGENESIS
; FILE REFERENCE: 27708/04013
; CURRENT APPLICATION NUMBER: US/09/881,578A
; CURRENT FILING DATE: 2001-06-14
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 2143
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-881-578A-3

Query Match 34.9%; Score 703; DB 3; Length 2143;
Best Local Similarity 69.7%; Pred. No. 1,2e-212;
Matches 1006; Conservative 0; Mismatches 410; Indels 27; Gaps 3;

QY 172 TTGAGTTCTTAAACAATGAGCCCAAGGTTTGTGAGCGTTGCCGAGCTGACCTTACCTTT 231
Db 10 TGCATTTCCAGAACCATGACTGATGGGTGTGAGCAATTAAGGAATGTGGCCATTCGACTTC 69
QY 232 TCTCAGAGAGAGGGGCTGTCTGAACCTGTCTCAGAGGACCTGTACTGGGACGTGATG 291
Db 70 TCTCAGAGAGAGGGGATGCTTGAACCTGTCTCAGAGGACCTGTACTGGGATGTGATG 129
QY 292 CTGAGAACTACAGTAATCTGTCTCAGATTTTGAATC--AGCATATGAAAAATAG 348
Db 130 TTGAGAACTATAGTAACTTGGGTCACTGGAATTTGAGATCAAAAAAGTATGAGACCAA 189
QY 349 AGTTTACTTACGAAAAAACAATCATGAAATTAAGGCTTCCAAAAGAAATTCAGTAA 408
Db 190 AAAATATTTTTCAGAAATATATATTTTGAATAAATTTTCCAGTGGAGATGAAGGAC 249
QY 409 AGAAGTAATCCCTTGAGCC-----GTAAGTATATGTGAAGT 447
Db 250 AAAAGTAATCCCTTGAGCC-----GTAAGTATATGTGAAGT 309

```

OY 448 AGCTTGAAGAACCAAGCGCTCCAGAGGAGGTATGTCATCAGATGATCATTTAT 507
DB 310 ATATTGAGGGACTTAAAGAGACATCAAGGGATCTTACGTCAATGATTAATCGCTAT 369
OY 508 GTCAAAAGGCGCTGCTACTAGAGAGGCAACCCCTCTAGAACATCATAGAGA---CATCAT 564
DB 370 GAAAAAATACCTTTCTTACAGAAAAAGTAAATCTTACTCCATCAATCAAGAAATTCATAT 429
OY 565 AAGGGAATTCCTTTGATGTATAGACTGTGGGAGGCGTTTATGTGTGGCTATCACTT 624
DB 430 ACAGAGAAATCCTATGTTGTATAGGAATGTGGGAGGCTTGCAGTCATGGCTCAAAACTT 489
OY 625 AGTCAACATCAGAAATCATCTGTGTAGAAACCTTATGAAATGTAAGATGTAAGAAG 684
DB 490 GTTCAACATGAGAAATCTATACAGCTGAAAGCACTTGAATGTAAAGAAATGGGAAG 549
OY 685 GCGTTCCGTGGGGCAATCAAGCTTACTCAACATCAAAAAATTCATACTGGGAGAGAGCC 744
DB 550 AATTATTTAAGTGCCTATCACTCAATGTGCATCAGATTTTCATCTGTGAGAAAGCC 609
OY 745 TACGATGTAAAGACTGTGGGAGGCTTTTCATGGGCGTCAAGCGCTGTTATTCATAG 804
DB 610 TATGAGTGAAGAAATGTGGGAGAGCTTTAGCTGGGATCAAGCGCTGTTAAACATGAG 669
OY 805 AGGATTCATCTGAGTGAAGAAACCTTATGAAATGTAAGAACTGTGGAAAGCGCTTGGCGT 864
DB 670 AGAATTCACACTGTGAGAGAACCTTATGATGTAAGAAATGTGGGAGGCGCTTATAGTGT 729
OY 865 GGTGATGAGCTCACTCAGACAGAGATTCACACTGGGAGAGAAAGACTACGATCAAA 924
DB 730 GGCCTACACTTACCAACATCAGAAATTCATATTGGTGTAAATCTTATTAAGTAAAG 789
OY 925 GACTGTGGGAGAGCTTTAGCCGTGTATTAACCTTATCAGCAAGAGAAATTCATAGT 984
DB 790 GAATGTGGGAGGCGCTTTTGTGGGCTCAAGCGCTGCTAAACATGAGATTAATTCATCA 849
OY 985 GGGGAGGAGCGCTTACGAGTGAAGACTGTGGGAGGCGCTTATTTGTGGTCAAGCGCTC 1044
DB 850 GGTGAGAAACCTTATTAATGTAAAGAAATGTGGGAGGCGCTTCACTGTGGCTATCACTT 909
OY 1045 ATTACAGATAAAAGAAATTCACAGAGTGAAGAAACCTTATGAAATGTCAAGAAATGTGGAG 1104
DB 910 ACTCAGATCAGAAATCATCTGTGTAAAGAAACCTTATGAAATGTAAATATGTGAAAG 969
OY 1105 GCGTTTACTCGAGTAAATTAATTAATCAAGCATCAGAGATCCACCGGTGAGAAAGCTT 1164
DB 970 GCTTTTGTGGGCTATCACTTACTGCAATCAGATATTTTCATATCTGTGAGAAAGCC 1029
OY 1165 CACGAATGAAGAGAGTGTGGGAGGCGTTTGGCTGGGCTTGAAGCGCTGTTAAGACGAG 1224
DB 1030 TATGATGCAAGAAATGTGGGAGGCGTTTAAATGGGATCAAGCTTATTCACATGAA 1089
OY 1225 AGGATACATACGCGGAGAGGCGCTACAGTGCAGAGATGTGGGAGGCGCTTCAATTGT 1284
DB 1090 AGAATTCATACGTGTGAGAGAACTTATGAAATGAAGATGTGAGAAAGCGCTTATAGTGT 1149
OY 1285 GCGTTCACCTCACTCAGACGAGAGATCCACAGCGGAAACCCCGTATTAATGTAG 1344
DB 1150 GGCCTATCACCTTTTCAACATCAGAAATCCATCTGTGAGAAAGCGCTTGAATGTAG 1209
OY 1345 GAGTGTGGGAGGCGTTTCACTTATGATGAGTGCCTGTGAAACATGAGAGATTCATACC 1404
DB 1210 GAATGTGGGAGGCGCTTATGTTGTGGGTTCAAGCGCTTGTAAACATGAGAGATTCATCT 1269
OY 1405 GGGGTGAAGCCCTATGGGTGTACAGAAATGTGGGAGGCGTTTATGTCACGCGCATCAGCTT 1464
DB 1270 GGTGAGAAATCCATGAATGTAAAGATTCGGAAGAGACTTTTGTATGTGGATCAAGCTT 1329
OY 1465 ACACAAATCAGAAAAAGCAGTGTGGGAGAAATCTTACGATGTAAAGAGTGTGGGAG 1524
DB 1330 ACTTCAACATCAGGTATTTTCACTGTGTGAGAAACCTTATGAATGTAAAGAAATGTGGGAG 1389
OY 1525 GCATGTAAACCACTTAACCATCTCGAGAAACATCAGAGGATCCACAAAGTTGAGAGAGCC 1584

```

```

DB 1390 GCTTTAATGTGATCAAGCTTGTTCACATGAAAGATTCATACAGGGAGAAAGCC 1449
OY 1585 TTT 1587
DB 1450 TAT 1452

RESULT 3
US-09-620-312D-154
; Sequence 154, Application US/09620312D
; Patent No. 6569662
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Feiyan
; APPLICANT: Chen, Rui-hong
; APPLICANT: Zhao, Qing A.
; APPLICANT: Wehrman, Tom
; APPLICANT: Xue, Aidong J.
; APPLICANT: Yang, Yonghong
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Zhou, Ping
; APPLICANT: Ma, Yuning
; APPLICANT: Wang, Dunrui
; APPLICANT: Wang, Zhiwei
; APPLICANT: John Tillinghast
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. 6569662e1 Nucleic Acids and
; FILE REFERENCE: Polypeptides
; CURRENT APPLICATION NUMBER: US/09/620,312D
; PRIOR FILING DATE: 2000-07-19
; PRIOR FILING DATE: 2000-04-25
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 1105
; SOFTWARE: pc_fl_genes Version 1.0
; SEQ ID NO 154
; LENGTH: 2274
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (116)..(460)
US-09-620-312D-154

Query Match 25.8%; Score 520.6; DB 3; Length 2274;
Best Local Similarity 63.1%; Pred. No. 1.4e-154;
Matches 930; Conservative 0; Mismatches 509; Indels 34; Gaps 7;

OY 121 CTCGGAATTTGCTTCTTCAAGTGAAGAACCCGAGAAAGCTGATCAGTTCTTCAAGTTCT 180
DB 56 CTCGGAATTTGCAACCTTCCAGAGAGAAAGCCGAAATTTGATCAAGTTC-----TAA 109
OY 181 AAAACAATGAGCCAGAGGTTTGTGAGAGTTCGCGGAGAGTGAAGCATAGACTTTTCTCAGAG 240
DB 110 AAAACAATGAGCCAGAGTTCGTTGTGTGACTTTCAAGGAGATGTGGCATTTGCTTCTCAGAG 169
OY 241 GAGTGGGCGCTGTCTGAATCTGTCTCAGAGGAGCTGTACTGAGAGCTGTATGCTGAGAGAC 300
DB 170 GAATGGGAATGTCTGAGACCTGTCTCAGAGGAGCTGTGTACAGGAGTGTATGTTGAGAGAC 229
OY 301 TACAGTAATCTGTGTCTCACTGGAATTTTGGAGTCAAGCATATGAAAAATGAAGTTTACTACA 360
DB 230 TACAGTAATCTGTGTCTCACTGGAATTTTGGAGTCAAGCATATGAAAAATGAAGTTTACTACA 289
OY 361 GAAAAAATTCATTCAGAAATTAAGGCGTTCCAAAGAAATTCAGATGAAGAGT----- 414
DB 290 GAAAAAATTCATTCAGAAATTAAGGCGTTCCAAAGAAATTCAGATGAAGAGT----- 349

```

```
QY 415 -----AAATCCCTTGGCCGTAACGTGATATGNGAGGTACGCTTGAAG 458
Db 350 CATCACCTTTCAATACATAGTGTCTTGGAGACAAATATGAGTGCAGAAAGGCACTTAGAGG 409
QY 459 ACCACAGCCCTCCAGAGGAGGTATGTCAATCATGATCATCAATTAATGTCAAAAGGCC 518
Db 410 TCAAGTAAGCAAGTCAAGAGGGGCTTACATGTGTGTCAAAATTAACCTGTGAAGAAAGGC 469
QY 519 TCGTACTAGAGAAAGGCACCCCTCTGTAAGAACATCACTGAGACATCAAT--AGAGAAATTC 576
Db 470 CACTGAAATCTATTCACCTCTTCTCACTTTTCATCGCAATTAATTCCTACCAAGGAAAAAT 529
QY 577 TTTGAATGTAGAGACTGTGGAGAGCCCTTATGTCGTGCTATCACTTATGTCAACATCAG 636
Db 530 GTCAAAATGTAGAGAAATGACAGACAGTTTCACTGCTCATGCTTATTAACATGAG 589
QY 637 AAAATCCATACGTGTGAGAAACCTTATGATGTAAAGAAATGTAAAGAGCCCTTCGTTGG 696
Db 590 GAAATCATATATATATATAAAATGCTCTGAAGTTAATTAACAAGAAATCCTTAGCAAA 649
QY 697 GGCATCAGCTTACTCAACATCAAAAAATTCATCTGGG--GAGAAGCCCTAGCAATGTAA 755
Db 650 AAGCCAAAGCTATATTTTAAACATCAGAAATTCAGACTGGGTGAGAAACCTTAGAGGTAT 709
QY 756 AGACTGTGGGAAAGGCTTTTCGATGGGGCTCAAGCCTCGTTATTCATTAAGAGATTCATAC 815
Db 710 GGAATGTGAAAGAGCCTTTGTGTGTCTTGTGATCTCATTAACATCAAGAAATTCATAC 769
QY 816 TGGTGAAGAAACCTTATGAATGTAAAGACTGTGAAAGCCCTTTCGGCGTGTGATGAGCT 875
Db 770 TATGAAAAACCTTATCAAGTGTAAAGCATGTGGAAAGCTTTTATTCGTGTGTACAGCT 829
QY 876 CACTCAGACACAGAGATTCACACTGGGAGAAAGACTAGATGCAAGAACTGTGGAA 935
Db 830 CACTGAACATCAAGAGGTTCAACAGAGAGAAACCAATGATGTGAAGAAATGTGAAA 889
QY 936 GACCTTTAGCCGCTGTATTAACCTTATTCAGACAAGAGAAATCATAGTGGGAGAGCC 995
Db 890 AGCTTTATGTAT--TGTCACAATATACTCTTCATCAGAGAAATCATAGTGTGAAAAAC 948
QY 996 TTAAGAGTGAAGAGCTGTGGAGAGCCTTATTTGTGTGTCAAGCCTCATTCAGCATAA 1055
Db 949 CTAATGATGT--AAGATTTGTGGAGAGCCTTATCT--TGGCTCTCAACTTACTTACATCA 1006
QY 1056 AAGAATTCACACAGGTGAGAAACCTTATGAAATGTCAAGATGTGGAGAGCCTTACTCG 1115
Db 1007 GAGAAATTCATAGTGTGAGAAACCTTATGAGTGAAGAAATGTGGAGAGCCTTATTTCT 1066
QY 1116 AGTCAATTAACCTTACTCAGCATCAGAAAGATCAACCGGTGAGAGAGCCTCAGCAATGTAA 1175
Db 1067 TGTGTACACCTTACATACATCAGAGAGTTCACTGTGTGAAAGCCTTACATATGTAA 1126
QY 1176 GAGGTGTGGAGAGCCTTTGCTGTGGGTTTCAAGCCTGTTAAGCAGAGAGATACATAC 1235
Db 1127 AGAATGTGGAGAGCCTTTTATGTGCTCCCAACTGAATGAACATCAGAGAAATTCATAC 1186
QY 1236 GGGCGAGAGAGCCTTCAAGTGCACAGAAATGTGGAGAGCCTTCAATTTGTGCTATACCT 1295
Db 1187 AGAGAGAGAAACCTTATGAAATGTAAAGATCGGGAGAAACCTTTTTCGTGTGCTCACACT 1246
QY 1296 CACTCAGACAGAGAGATTCACACAGGAGAAACCCGTATTAATGTAAAGAGTGTGGAA 1355
Db 1247 TACTTACACCTGAGAGTCTTACGTGAGAGAGCCTTATTAATGTAAAGATGTGGAA 1306
QY 1356 GGGCTTCATTTATGATGAGAGCCTCGTGAACATGAGAGATTCATACCGGGGTGAAGCC 1415
Db 1307 AGCCTTATTTCTAATCTTATCTTATTCACATCAAGAAATTCATACCGGAGAGAGCC 1366
QY 1416 CTATGGGTGTACAGATGTGGAGAGAGCTTTAGTACAGGCAATGAGCTTACACAACTCA 1475
Db 1367 CTACAAATGTAAAGAAATGTGAAAGAGCCTTTATTTGTGCAAAACATCTAGTGAACATCA 1426
```

```
QY 1476 GAAAAAGCAAGTGTGGGAGGAAATCTTACGAATGTAAAGATGTGGAGAGCATGTAAACA 1535
Db 1427 GAGAAATTCATACAGGTGAGAAACCTTTGAATGTAAAGAAATGTGAAAGAGCCTTATTCG 1486
QY 1536 CCTAAACCATCTCCGAGAAACATCAGAGGATCCA 1568
Db 1487 TGTTCATATCTTACTCAACATGAGAAAAATTCGA 1519

RESULT 4
US-09-774-528-331
; Sequence 331, Application US/09774528
; Patent No. 6743619
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Zhou, Ping
; APPLICANT: Goodrich, Kyle
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Ren, Feiyan
; APPLICANT: Zhang, Jie
; APPLICANT: Zhao, Qing A.
; APPLICANT: Yang, Yonghong
; APPLICANT: Xue, Aidong J.
; APPLICANT: Wehrman, Tom
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Wang, Dunrui
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. 6743619el Nucleic Acids and
; FILE REFERENCE: 802
; CURRENT APPLICATION NUMBER: US/09/774,528
; CURRENT FILING DATE: 2001-01-30
; NUMBER OF SEQ ID NOS: 441
; SOFTWARE: pc_Fl_genes Version 2.0
; SEQ ID NO 331
; LENGTH: 2499
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (222)..(1820)
; US-09-774-528-331

Query Match 25.3%; Score 510; DB 3; Length 2499;
Best Local Similarity 63.6%; Pred. No. 3 6e-151;
Matches 821; Conservative 0; Mismatches 445; Indels 24; Gaps 2;

QY 322 GATTGGAGTCAGCATATGAAAAATAAGAGTTTACTTACAGAAAAAAACATTCATGAATA 381
Db 450 GATTGGAGCCAAAGTATGAGACCAAGAAATTATCTTTGAAAAATGACATTTATGAATA 509
QY 382 AGGGCTTCCAAAAGAAATTCAGATAGAAAGTAATCCCTTGGCCGTAAAC----- 432
Db 510 AATTATCCCAAGTGAAGTAATGAAAGAAATGAAACCAATGAGCCTTAAAGGTCTCATT 569
QY 433 -----TGAATGTGAAGGTACGCTTGAAGAAGCAACAGGCTCCAGAGGAGG 480
Db 570 TTAAGAAATGATTTGGAAATCCACAGAAATTTGAAGACAGAGAGAGACTCAAGAGGA 629
QY 481 TATGTCAATCAGATGATCATTAATGTCAAAAGGCTGTCTACTAGAGAGGACCCCT 540
Db 630 TACTTAGTAGTGTGAATAGCCATCTGAAGAGGTCTCTTACAGAAACGACATCT 689
QY 541 CTTGAAACATATAGAGATCA--TAAAGAAATTCCTTGAATGTAAAGACTGTGG 597
Db 690 GTTACTCACAATCAGAGACTTCAATTTGTGATTAACCTTATGAATGTAAAGATGTGG 749
QY 598 AAGGCTTTAGTGTGCTATCAACTTAAGTCAACATCAGAAATTCATATGCTGTGAGAA 657
Db 750 AAGGCTTCAAGATGTGGCCCAAGCTTACTTTTCAATCAGAAATTCATATGCTGTGAGAA 809
QY 658 CTTATGAATGTAAAGAAATGTAAAGAGCCTTCCGTGGGCAATCAGCTTACTCAACAT 717
```

```

Db      810 CCGTATGATGTAAAGAAATGTGGATGGCTTTCAGACAGACTGCACACTTACTGCACAT 869
Qy      718 CAAAAAATTCATATCTGGGAGAAAGCCCTTACGAATGTAAAGACTGTGGAAAGCCCTTTCGA 777
Db      870 CAGAGACTTCATCTGTGTGTAAGAACTCTATGATGTAAAGAAATGTGGGAAAGCTTTCATA 929
Qy      778 TGGGGCTCAAGCTCTGTTATTCATTAAGAGATTCATCTGTGTGTAAGAAAGCCCTTATGATGT 837
Db      930 TGTGTGTGCAAGATCTTAAGATACATCAAGAAATGCAATTTGTGTGTAAGAAAGCCCTATGATGT 989
Qy      838 AAAAGCTGTGAAAGGCCCTTTCGGCGGTGTGTGATGAGCTCACTCAGACACAGATTCAC 897
Db      990 AAAAGATGTGGAGAGCCCTTTCAGGATGAGAGACAACTTACTCTGCATGAGAGATTCAT 1049
Qy      898 ACTGGGAGAAAGAACTTACGAATGCAAGAATGTGGAGAACCTTTAGCCGTGTGTATATA 957
Db      1050 ACTGTGTGAGAAACCTTATGTGTGTAAAGAGTGTGGAAAGCCCTTTAGACATGACGACAC 1109
Qy      958 CTTATTCAGCACAAAGAAATTCATATGTGGGAGAAAGCCCTTACGAGTGTAAAGACTGTGGG 1017
Db      1110 CTGACTCGGCATCAGAAAGCTTATATGTGTGACAGGCTCTATGAAATGCAAGAAATGTGGG 1169
Qy      1018 AAGGCTTTATTTGTGTGTCAAGCCCTCATTCAGCAATAAAGAAATTCACACAGGTAGAAA 1077
Db      1170 AAGGCTTTTGTGTGTGTCTGTGTCTTATGATGATCAACAACTTCATCTGTGTGAGAAA 1229
Qy      1078 CCTATGATGTCAAGAAATGTGGAGAGCCCTTTCAGTCAATTCCTTACTCGACAT 1137
Db      1230 CCTATGATGTAAAGAAATGTGGAGAGCCCTTTCAGTGTGCAACAACTTCATCTGTGTGAGAAA 1289
Qy      1138 CAGAAATTCACACCGGTGAGAAAGCCCTCAAGATGTAAAGAGTGTGGAAAGCCCTTTCGC 1197
Db      1230 CAGAAATTCACATCTGTGTGAGAAAGCCCTTATGATGTAAAGAAATGTGGAAAGCCCTTTCGC 1349
Qy      1198 TGGGGTGTGAGCCGTGTAAAGACGAGAGATTAATACGGGGGAGAAAGCCGTAAAGTGC 1257
Db      1350 CTTGTGTATCATCTTATCTTCATCAACAAATTCATCTGTGTGAGAAAGCCCTTATGATGT 1409
Qy      1258 ACAGAAATGTGGAGAGCCCTTTCATTTGTGTGTATCACTCTCAGACAGAGAAATTCAC 1317
Db      1410 AAGGAATGTGTAAAGCCCTTATGTGTGTATCTCACAACCTTATTCATCTGTGTGAGAAA 1469
Qy      1318 ACAGGCGAAACCCCTTATTAATGTAAAGAGTGTGGAAAGCCCTTTCATTTATGATGTGAGC 1377
Db      1470 ATTTGTGTAAAGCCCTTATGTATGTAAAGATGTGGGAAAGCCCTTACAGACTTTCACAA 1529
Qy      1378 CTTGTGAAACATGAGAGATTCATACCGGGGTGAAACCTTATGGGTGTACAGAAATGTGGG 1437
Db      1530 CTCACACAGCATCAAGATGTATCTATTTGTGTGAGAAACCTTATTAATGTAAAGAAATGTGGC 1589
Qy      1438 AAGACTTTAGTCAAGGCGCATGAGCTTACACAACATCAGAAAGAGCAGAGTGGGCGAAA 1497
Db      1590 AAGGCTTTTATGATGTGGCGCAAAAACCTTACTACATCAGAGCACTTATCTGTGGCGAAAAA 1649
Qy      1498 TCTTACGAATGTAAAGAGTGTGGGAAAGGATGTAAACCACTTAAACCATTCGGAAGAAC 1557
Db      1650 CCGTTTGTGTGTAAAGAAATGTAGAAAGCCCTTTCAGCTTAAATTCATCTCTTATTCACAT 1709
Qy      1558 CAGAGATTCACACAGAGTGTGAAGAGCTTT 1587
Db      1710 CTGAGAAATTCATCTGTGTGAGAAACCTTAT 1739

```

RESULT 5
 US-10-120-988-331
 ; Sequence 331, Application US/10120988
 ; Patent No. 6919193

; GENERAL INFORMATION:
 ; APPLICANT: Tang, Y. Tom
 ; APPLICANT: Goodrich, Ryle
 ; APPLICANT: Liu, Chenshua
 ; APPLICANT: Ren, Feiyun

```

; APPLICANT: Wang, Dunrui
; APPLICANT: Drenmac, Radoje T.
; TITLE OF INVENTION: No. 6919193el Nucleic Acids and
; TITLE OF INVENTION: Polypeptides
; FILE REFERENCE: 802CON
; CURRENT APPLICATION NUMBER: US/10/120,988
; PRIOR FILING DATE: 2002-04-11
; PRIOR FILING DATE: 2001-01-30
; NUMBER OF SEQ ID NOS: 441
; SOFTWARE: pt FL_genes Version 2.0
; SEQ ID NO 331
; LENGTH: 2499
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (222)..(1820)
US-10-120-988-331

Query Match      25.3%; Score 510; DB 3; Length 2499;
Best Local Similarity 63.6%; Pred. No. 3.6e-151;
Matches 821; Conservative 0; Mismatches 445; Indels 24; Gaps 2;

Qy      322 GATTGTGAGTCAGCATATGAAGAAATTAAGACTTACTTACAGAAAGAAACATTCATGAATA 381
Db      450 GATTTGGAGACCCAGATATGAGACCAAGAAATTTATGAGAAATGACATTTATGAAATA 509
Qy      382 AGGCTTCGAAAGAAATTCAGATGAGAAAGTAATCCCTTGGCGGTAAAC----- 432
Db      510 AATTTATCCAGTGTGAAGATTAAGAAAGATTTGAAACCAATGGCTTTAAGGCTTCAT 569
Qy      433 -----TGAATGTGAAGTACGCTTGAAGACCAAGCGCTCAGAGGAGG 480
Db      570 TTAAGAAATGATTTGGAAATCCACAGAAATTTGAAGACAGAGAGACCTCAAGAAAGA 629
Qy      481 TATGTCAATCAGATATATCATATATGTCAAAAGCGCTGTCTACTAGAGAAAGCACCT 540
Db      630 TACTTCAGTGTGTAAAGATGCAATCTGAAAGGTGTCTCTTACCAAGAAAGCACATCT 689
Qy      541 CCTAGAACATCAGAGCATCA---TAAGGAAATTCCTTGAATGTAAAGACTGTGGG 597
Db      690 GTTACTCCAGATCAGAGACTTATTTGTGTATTAACCTTATGATGTAAAGAAATGTGGG 749
Qy      598 AAGGCTTTAGTGTGTGTATCACTTATGATCAATCAGAAATTCATATCTGTGTGAGAAA 657
Db      750 AAGGCGTTCAGAGTGTGCGCAACAGACTTCTTTCATCAAGAAATTCATCTGTGTGAGAAA 809
Qy      658 CCTATGAAATGTAAAGATGTAAAGAGCCCTTCCGTGGGCAATCAGCTTACTCAACT 717
Db      810 CCGTATGATGTAAAGAAATGTGGAGTGGCTTTCAGACAGACTGACACCTTACTGCACAT 869
Qy      718 CAAAAAATTCATATCTGGGAGAAAGCCCTTACGAATGTAAAGACTGTGGAAAGCCCTTTCGA 777
Db      870 CAGAGACTTCATCTGTGTGTAAGAAACCTTATGATGTAAAGAAATGTGGGAAAGCTTTCATA 929
Qy      778 TGGGGCTCAAGCTCTGTTATTCATTAAGAGATTCATCTGTGTGTAAGAAAGCCCTATGATGT 837
Db      930 TGTGTGTGCAAGATCTTAAGATACATCAAGAAATGCAATTTGTGTGTAAGAAAGCCCTATGATGT 989
Qy      838 AAAAGCTGTGAAAGGCCCTTTCAGGCGGTGTGTGATGAGCTCACTCAGACACAGATTCAC 897
Db      990 AAAAGATGTGGAGAGCCCTTTCAGGATGAGAGACAACTTACTCTGCATCAGAGATTCAT 1049
Qy      898 ACTGGGAGAAAGAACTTACGAATGCAAGAATGTGGAGAAAGCCCTTTCAGCGGTGTGTATATA 957
Db      1050 ACTGTGTGAGAAACCTTATGTGTGTAAAGAGTGTGGAAAGCCCTTTAGACATGACGACAC 1109
Qy      958 CTTATTCAGCACAAAGAAATTCATATGTGGGAGAAAGCCCTTACGAGTGTAAAGACTGTGGG 1017
Db      1110 CTGACTCGGCATCAGAAAGCTTATATGTGTGTGACAGGCTCTATGATGTCAAGAAATGTGGG 1169
Qy      1018 AAGGCTTTTATTTGTGTGTCAAGCCCTCATTCAGCATAAAGAAATTCACACAGGTGAGAAA 1077

```

Db	1170	AAGGCGCTTTTGTGTGGCTCTGTGGTCTTAGAGTACATCAAACTTCACTACTGGTGGAAA	12229
QY	1078	CCCTATGAATGTCAAGAAATGTGGGAAGGCTCTTTA-CTCGAGTCAATTACTTACTCAGCAT	11377
Db	1230	CCCTATGAATGTGAAGAAATGCGGGGAAGGCTTTTAGAGTGCACAACAACCTAACCTCCAT	12899
QY	1138	CAGAAGATCCACACCGGTGAGAAGCTTCACGAATGTAGAGATGTGGGAAGGCTTTGCG	11977
Db	1290	CAGGAATTCATCTACTGGTGAAGAAACCTTATGAATGTAGGAATGTGGAAAAGACTTTTAGC	13499
QY	1198	TGGGGTTTCGAGCCTCGTTTAAAGCAGAGAGATACATACGGGCGAAGAGCCGTACAAATGC	12577
Db	1350	CGTGGTTATCATCTTATATCTTCATCACAAGAAATTCAATCTGGTGAAGAAACCTTACGAATGT	14099
QY	1258	ACAGAAATGTGGGAAGGCTTTCAATTGTGGCTTATCACTCACTCAGACAGAGAAATCCAC	13177
Db	1410	AAGGAATGCTGGAAAAGCCTTTTAGTGTGCTACTACACAACTTATTTTCAATCAAGAGATTCAT	14697
QY	1318	ACAGGCGGAAAACCCGATATAATGTATGAAGAGTGTGGGAAGGCTTTTCATTTATGATTCGAGC	13777
Db	1470	ATTGTGTTAAGCCCTTATGACTGTATAGGAATATCGGGGAAGGCTTACAGCTACTTTCACAA	15297
QY	1378	CTCGTGAACAATGAGAAATTCATACCGGGGTGAACCTTATGGGTGTACAGATGTGGG	14377
Db	1530	CTCACACAGCATTCAGAGATTCATATTGGTGAAGAAACCTTATAATGTATGAAGAAATGTGCG	15897
QY	1438	AAGAGCTTTAATGACAGGCGCATCAGCTTACACAAACATCAGAAAAACGACAGTGGGCGAAA	14977
Db	1590	AAGGCGCTTTAATGTGGCGCCAAAACTTACTTACATCAGAGCATTCATCTGGGGAAGAAA	16497
QY	1498	TCCTACGAATGTAAAGAGTGTGGGGGAAGGCATGTAAACAACCTTAATCTCCGAGACAT	15577
Db	1650	CCCTTTGATGTATGAAGAAATGTAGGAAGGCGCTTTAAGACTTAATTCATCCCTTATTCACAT	17097
QY	1558	CAGAGGATCCACAACAGTTGAAGAGCCTTT	1587
Db	1710	CTGAGAAATTCATTTCTGTGTGAGAAACCTTAT	1739
RESULT 6			
US-10-104-047-415			
Sequence 415: Application US/10104047			
Patent No. 6943241			
GENERAL INFORMATION:			
APPLICANT: HELIX RESEARCH INSTITUTE			
TITLE OF INVENTION: No. 6943241e1 full1 length cdna			
FILE REFERENCE: H1-A0105			
CURRENT APPLICATION NUMBER: US/10/104,047			
CURRENT FILING DATE: 2002-03-25			
PRIOR APPLICATION NUMBER:			
PRIOR FILING DATE:			
NUMBER OF SEQ ID NOS: 4096			
SOFTWARE: PatentIn Ver. 2.1			
SEQ ID NO 415			
LENGTH: 3472			
TYPE: DNA			
ORGANISM: Homo sapiens			
US-10-104-047-415			

Query Match	Similarity	Score	502.8	DB	3	Length	3472
Best Local	Similarity	67.0%	Pred.	8.8e-149			
Matches	714	Conservative	0	Mismatches	352	Indels	0
OY	559	CATCATTAAGAGAAATTCCTTTGAATGTAAGGACTGTGGAGAGCCCTTTAGTCGTGGCTAT	618				
Db	769	CATCTGGAGAGAAACCTTTTGAATGTAAGGACTGTGGAGAGCCCTTTGCAATTCACATA	828				
OY	619	CAACTTAGTCAACATCAGAAAATCCATACCTGTGTGAGAAACCTTATGATGTAAGAAATGT	678				
Db	829	CAATTACTCGACATCAGAAATTTTCACTGGTGAAGAAACCTTTTGAATGTAACGAATGT	888				
OY	679	AAGAGGCCCTTCGTTGGGCAATCAGCTTACTCAACATCAAAAATTCATACCTGGGAG	738				

Db	889	GGAAAGGCGTTTAACTCTTCTTAACCTGCTTAATGCGATGAAGAACATTCAACAGGCGAG	948
Qy	739	AAGCCTTACGATGTAAAGACGTGGGAAGGCTTTTGCATGGGGCTCAAGCTTGTTATT	798
Db	949	AAACTGTTGAATGTAAAGAAATGGGGAAGCTCTTTAATGTAAGCTCAAACTTGTTCAA	1008
Qy	799	CATAAGAGATTCATCTCGGTGAAAAAACCTATGAATGTAAAGCTGTGAAAGGCTTT	858
Db	1009	CATCAGATATTCTTCCTGGTAAACCATATGAAATGTAAAGAGTGTGGAAAGGCTTT	1068
Qy	859	CGGCGGTGTATGAGCTCACTCAGACACAGAGATTCACACTGGGGAAGAAAGCTACAA	918
Db	1069	AATGTGGTGCACCTTATTACACATCAGAAATTCATTCCAATGAGAAACCTTTGTA	1128
Qy	919	TGCAAAAGCTGTGGAAAGACCTTTAGCCGTGTATATACTTATTCAGCACAGAGATT	978
Db	1129	TGTAAAGAAATGTGGGAATGGCCCTTTCGATTCATTACCACTTATTTGAACATTCGCAATT	1188
Qy	979	CATATGGGGGAGAAAGCTTCAGAGTGTAAAGACTGTGGGAAGGCTTTTATTTGTGTCTCA	1038
Db	1189	CATACTGTGTGAAGAACCTTTTGATGTAAAGATGTGAAAGGGCTTACTCTTCAGCA	1248
Qy	1039	AGCCTCATTCAGCATAAAGAAATTCACACAGGTGAGAAACCTATGAATGTCAAGAAATGT	1098
Db	1249	AAGCTGTTGCACATCAGAGAAATTCATACGTGTGAGAAACCTTTGAATGACAGGGAATGT	1308
Qy	1099	GGGAAGGCTTTACTCGAGTCAATTACTCTTACTCAGATCAGAGATCCACACCGGTGAG	1158
Db	1309	GGGAAGGCTTTAGCTTCTTCAACAGCTTATGTCCATTAAGAACATTTACACAGGTGA	1368
Qy	1159	AAGCCTCAGATGTAAAGAGTGTGGGAAGGCTTTGCTGGGGTTCGAGCCTTGTTAAG	1218
Db	1369	AAACCGTTTGATGTAAAGAGATGTGGGAAGTCTTTAATGTACTCAAACTTGTTCAA	1428
Qy	1219	CACGAGAGATPACATACGGGCGAGAGGCCGTTCAGATGTGCACAGATGTGGAGAGCCTTC	1278
Db	1429	CATCAGAGTATTCTAGCTGTATAAACAATATCAATGTAAGAGTGTGGAAAGGCTTT	1488
Qy	1279	AATTGTGCTATCACTTCACACACGAGAGATCCACACAGGCGAAACCCGTAATPAA	1338
Db	1489	AATGTGTGTGACACCTTATTCAGATCAGAAATTCATTCCAATGAGAAACCTTTGTGA	1548
Qy	1339	TGTAAAGAGTGTGGGAAGGCTTTGATTATATGATCGAGCCTCGTGAACATGAGAGATT	1398
Db	1549	TGTAGGGAATGTAGATGTGCTTTAATGATATCATTGCCAATTAATTAAGCAATTCGAATT	1608
Qy	1399	CATACCGGGGTGAACCTTATGGGTGTACAGATGTGTGGGAAGAGCTTTATGTCAAGGCCAT	1458
Db	1609	CATATCGTGTGCAAGCCCAATTGTAATGTCAAGACTGTGGGAAGGCCCTTCAATCGTGCTCA	1668
Qy	1459	CAGCTTACACAACATCAGAAAAAGCAGCTGTGGGCGAAATCCCTACGAATGTAAAGAGTGC	1518
Db	1669	AGCTTTGTTCACATCAGAGTATTCACACTGTGTGAGAGCCCTTATGAATGTAAAGAGTGT	1728
Qy	1519	GGGAAGCATGTAAACCACTAAACCATCTCCGAGAACATCAGAGAGATCACAACAGTTGA	1578
Db	1729	GGGAAGGCTTTTAACTTTACTTACCTCAACTTTCCCAACATCAGAAAACTCACACAGGTGAA	1788
Qy	1579	AGAGCTTTTGAACGCAAGTAGCCGCTGTATCTATAGTTTCGCTT	1624
Db	1789	AAACCATTTGAATGTAAAGAAATGTGGAAATTTCTTTCGTGTGGTT	1834

```

RESULT 7
US-10-104-047-1500
; Sequence 1500: Application US/101040407
; Patent No. 6943241
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: NO. 6943241el full length cDNA
; FILE REFERENCE: H1-A0105
; CURRENT APPLICATION NUMBER: US/10/104.047

```

```

; CURRENT FILING DATE: 2002-03-25
; PRIOR APPLICATION NUMBER:
; PRIOR FILING DATE:
; NUMBER OF SEQ ID NOS: 4096
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1500
; LENGTH: 1694
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-104-047-1500

Query Match      24.6%; Score 495.4; DB 3; Length 1694;
Best Local Similarity 61.5%; Pred. No. 1,3e-146;
Matches 878; Conservative 0; Mismatches 516; Indels 33; Gaps 4;

QY 182 AAACAAATGCCCCAGGGTTGGTGTGACGTTGCGCCAGCTGAGCCATGACATTTTCTCAGAGAG 241
DB 254 AAGCCATGGCTGAGGGATCAGTGTGATGTTCAAGTATGTGTCATAGACTTCTCTCAGAGAG 313
QY 242 AGTGGGCTGTCTGAACTGTCTCAGAGAGACTGTACTCTGGAGCGTGATGCTGGAGAACT 301
DB 314 AGTGGGACTGCTGTGACCTGTCTCAGAGAGACTGTATACAGAGATGTGATGTTGGAGAACT 373
QY 302 ACAGTAACCTGGTCTCACTGG-----ATTGAGAGTCAGCATATGAAATAGAGTTT 353
DB 374 ACGGCAATCTGGTTTCAATGGAGCTTTACACTCTTAAGCTTAAGTATCTCTTATTGG 433
QY 354 ACCTACAGAAAAAATTCATCATGAAATAAGGGCTTCCAAAGAAATTCAGATAGAGAG 413
DB 434 AACAAAGGAAAGAGCCCTGATGTTGGCAGAGACTTACAAAGAGCCGTGTTCAGATC 493
QY 414 TAAATCCCTTGCCCTTAAC-----TGAATATGTGAAGGTACCTTGAAAGACCAAG 465
DB 494 TGAATTCATGTGTGAAACCAAGTTATATCTTAAAGAGAAAGATTATGAAATAGAT 553
QY 466 CGCTCAGAGGGAGGTA-----TGTCATCAGATGATATCATTAATGTCAAAAGGC 517
DB 554 TATGCCAAGGGAGTAATGAGCACTTACAAAGCAAGCCCTTGAATCTCCAGTTTGGAG 613
QY 518 CTGCTACTAGAGAGAGCAACCCCTCTAGAACATCAGAGCATCATTAAGAGAAATTCCT 577
DB 614 ATGTTTGGAAATATGAAAGCCACTTGGAAACAACTGGGATATCAATGGGCAATTTTA 673
QY 578 TTGAA-----TGTAAGACTGTGGAGAGCCCTTATGCTGTGCTATCACTTATGTC 628
DB 674 GTCAAGAAATATTCACCTCGTAATCATGCCCCAATTTATTCACAGACATTCCTTACTC 733
QY 629 AACATCAGAAAAATCCATCTGGTGAAGAACTTATGAATGTAAAGATGTAAAGAGGCCT 688
DB 734 TCCATCAAAATATTAATATGAAAGCAAGACCTTATGATGTAAAGAAATGTGAAAGGCCT 793
QY 689 TCCGTTGGGGCAATCAGCTTACTCAATCAAAATTCATATCTGGGAGAGCCCTTACG 748
DB 794 TTAGTCAGAACTCACAATTTTATTAACATCAGAAATTCATATGTGTAAGAAATCTTATG 853
QY 749 AATGTAAAGACTGTGGAGAGGCTTTTCATGATGGGCTCAAGCTCCTTATTCATTAAGAGA 808
DB 854 AATGTAAAGAGTGTGGAAATCTTTAGTGTGTTCATGTCATGTCGGCATCTGAAAAA 913
QY 809 TTCACTACTGGTGAAGAAACCTTATGATGTAAAGACGTGGAAGAGCCCTTGGCGGTGGTG 868
DB 914 TTCACTACTGGCGAAGAAACCTTGTAAATGTAAAGAAATGTGGAAGAGCCCTTCATGTGACT 973
QY 869 ATGAGCTCACTGAGCAGCAGAGATTCACACTGGGAGAGAAAGACTATCGAATGCAAAAGCT 928
DB 974 CATACCTTTCTCAATCATCAGAGAAATCCATACCGGTAAAGAAACCTATGAAATGTAAAGAT 1033
QY 929 GTGGGAAAGACCTTTAGCCGTGTGTATTAATCTTATCAGAGCAAGAAATTCATATGTGGG 988
DB 1034 GTGGGAAAGCCCTTATGTATGTCTCAAACTTATTTGACATCAGGAAATTCACACTGGG 1093
QY 989 AGAAGCCTTACAGAGTAAAGACTGTGGAGAGCCTTTTATTTGTGTTCAAGCCTCATTC 1048

```

```

DB 1094 AAAAACTTATGAATGTAAAGTATGTGGAGAAAGCCTTACTAAGAGCTCACAACTTTTTC 1153
QY 1049 AGCTTAAAGAAATTCACACAGTGAAGAAACCTTATGAATGTCAAGAAATGTGGAGGCCT 1108
DB 1154 AGCATGCAAGAAATTCATCAGGTGAGAAACCTTATGAATGTAAAGAAATGTGGCAAGCCT 1213
QY 1109 TTACTGAGTCAATTAACCTTACTGAGCATCAGAGATCCACACCGGTGAGAAAGCCTCAAG 1168
DB 1214 TTACCAAGCTCAAAAGCTTTGTTCAACTCAGAGAAATTCATCTGTGTGAGAAACCTTATG 1273
QY 1169 AATGTAAAGAGTGTGGAGAGCCCTTTCGCTGGGTTTCAGGCTCTGTAAAGCAGAGAGA 1228
DB 1274 AGTGAAGAGAAATGTGGCAAGACCTTTAGTATGAGTGTGACACTTATCATATCAGAGAA 1333
QY 1229 TACATACGGGCGAGAGAGCCGTACAAATGTCAAGAAATGTGGAGAGCCCTTCAATTTGGCT 1288
DB 1334 TTCACTACTGTGAGAGAAACCTTATGATGTAAAGAAATGTGGAGAGCCCTTTACTCAGAGCT 1393
QY 1289 ATCACTCACTCAGCAGAGAGAAATCCACACAGGCGAAACCCGTATTAATGTAAAGAGT 1348
DB 1394 CACAGCTTCGTACATCAGAGAAATTCACGCTGTGAGAGAAACCTTTGAATGTCTGAAT 1453
QY 1349 GTGGAGAGCCTTTCATTTATGATCGAGCTCTGTGAAAATGAGAAATTCATACCGGGG 1408
DB 1454 GTGGGAGAGCCTTTATCTAGAACTCAAACTTTTCACAGATCAGAGAAATTCATACAGATG 1513
QY 1409 TGAACCTTATGGGTATCAAGATGTGGAGAGAGCTTTATGTCACGGCCATCAGCTTACAC 1468
DB 1514 AAAAACTATGAATGTATGATGATGTGAGAAAGCCCTTAAATGAATGCTCAAACTTACTC 1573
QY 1469 AACATCAAGAAACGACGAGTGGGCGAAATCTCGATGATGTAAAGAGTGGCGGAGAGGAT 1528
DB 1574 GACATCTGAGAAATTCACCTGTGAAAGCCCTTATACCTGTAAAGAAATGTGGAGAGCTT 1633
QY 1529 GTAACCACTTAACCATCTCCGAGAACATCAGAGAGATCCACAACTG 1575
DB 1634 TTAGTAGTGGCTCGGATCTCATTCGTATCAGGGAATTCATATCTAT 1680

RESULT 8
US-09-620-312D-163
; Sequence 163, Application US/09620312D
; Patent No. 6569662
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Feiyan
; APPLICANT: Chen, Rui-hong
; APPLICANT: Zhao, Qing A.
; APPLICANT: Wehman, Tom
; APPLICANT: Xue, Aidong J.
; APPLICANT: Yang, Yonghong
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Zhou, Ping
; APPLICANT: Ma, Yundong
; APPLICANT: Wang, Dunrui
; APPLICANT: Wang, Zhiwei
; APPLICANT: John Tillinghast
; APPLICANT: Drenth, Radoje T.
; TITLE OF INVENTION: No. 6569662el Nucleic Acids and
; FILE OF INVENTION: Polypeptides
; FILE REFERENCE: 784CIP2B
; CURRENT APPLICATION NUMBER: US/09/620,312D
; CURRENT FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488,725
; NUMBER OF SEQ ID NOS: 1105
; SOFTWARE: pc_FL_genes Version 1.0
; SEQ ID NO 163

```


LENGTH: 2925
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (178) ..(2433)
US-09-620-312D-163

Query Match 23.7%; Score 478.2; DB 3; Length 2925;
Best Local Similarity 66.9%; Pred. No. 5.3e-141;
Matches 677; Conservative 0; Mismatches 335; Indels 0; Gaps 0;

559 CATCATTAAGAGAAATTCCTTGAATGAAGAGCTGTGGAGAGCCCTTATGCTGTGCTAT 618
1159 CATACTGGGAGAGGACCATACCTTTAATGAGTGTGAAAAGCCCTTATGCGCAGAGAGCC 1218
619 CAACCTTATGCAACATGAGAAATCCCTACTGTGTGAGAAACCTTATGAAATGAAGATGT 678
1219 CACTTATGGAACATCAGAAATTCATACGGGAGAAACCTTTAAATGTGATGATGT 1278
679 AAGAGAGCCCTTCGCTGGGCAATCAGCTTACATCAATCAAAAATTCATCTGCGAG 738
1279 GATTAACCTTACCGCAGAGACACACTTACTCAATCAAAAATTCATCTGAGAA 1338
739 AAGCCCTACGAATGTAAGACTGTGTGAAAGGCTTTTCATGTGGGCTCAAGCCTCGTAT 798
1339 AAAACCTATTAATGTATGAAATGTGAAAAGCCTTCAAGGGCCCTCACTTTATCCGT 1398
799 CATTAAGAGATTCATCTGTGAAAACCTTATGATGTAAGACTGTGTGAAAAGCCCTT 858
1399 CATCTATATATTCATCTGTGTGAAAACCTTATGATGTAAGACTGTGTGAAAAGCCCTT 1458
859 CGGCGTGTGATGAGCTCACTAGCAGACAGAGATTCACACTGTGGGAGAAAGCTACGAA 918
1459 AGCCAGACTCAAACTCTCACTAGCAGATCAAAAATTCATCTGTGGAGAAAGCCCTTAT 1518
919 TSCAAAGACTGTGTGGAAGACCTTACCGTGTATTAACCTTATTCAGCAAGAGAAAT 978
1519 TGTGCAATGTGTGAAATCTTTATGTTACTGTGTATCCCTCTCAACCGTGAAT 1578
979 CATAGTGGGAGAGAGCTTACGAGTGTAAAGACTGTGTGAAAGCCCTTATTTGTGTTCA 1038
1579 CATACTGAGAGAAACCTTACAAATGTAATGTGTGAAAGCCCTTACGTTACTCTCA 1638
1039 AGCCCTATCAGAGTAAAGAAATTCACAGGTGAGAAACCTTATGATGTCAAGATGT 1098
1639 TCCCTTATCAATGAGAAATTCACAGAGAGAAAGCCCTTATGATGTCAAGATGT 1698
1099 GGGAGGCTTTACTGAGTCAATTAACCTTACAGCATCAAGAGATCCACCGGTGAG 1158
1699 GGAAGGCTTTCAGTTATCTCTCAAGCTTATCAGCATCAAGAACTCACTCAAGAG 1758
1159 AAGCCCTACGAATGTGAAGAGTGTGGAAAGCCCTTTCGCTGGGTTTCAGCCTCGTTAAG 1218
1759 AAGCTTATGAATGTGAAGAGTGTGGAAAGCTTTTATTCGAGGTTCTCTTGTGAAG 1818
1219 CACGAGAGATACATCGGGCGAGAAAGCGTACAGTGCACAGATGTGTGGAAAGCCCTT 1278
1819 CATGAAGAAATTCATCTGTGAGAGAAACCTTATCAGTGTATGAATGTGTGGAAAGCCCTT 1878
1279 AATTGTGCTATACCTCACTCAGCAGAGAGATCCACAGCGAAACCCCGTATATA 1338
1879 AGTTATGTTCATCCCTTATTCAGCATATGGAAGATCCATCTGGAAGAACGACCTTCAAG 1938
1339 TGTAAAGAGTGTGGAGAGATTCACAGACATACACCTTACACGCTTAAGAAAT 1998
1939 TGTAAAGAGTGTGGAGAGATTCACAGACATACACCTTACACGCTTAAGAAAT 1998
1399 CATACGGGGTGAAGCCCTTATGTGTGTAAGAGTGTGGAAGGCTTTATGTCACGCGAT 1458
1999 CATACGGAGCCAGGCTTATGTGTGTCTGAGTGTGTGTAAGGCTTTTCAGACTTGTTC 2058
1459 CAGCTTACACAAATCAGAAAGCAAGTGTGGCGAATCTTACGAATGTAAAGAGTGC 1518

2059 TCTCTTGCTCAACATCAAAAACTCACAAGAAAGAAACCTTACAGGTATTAATGT 2118
1519 GGGAGGAGATGTAAACCATTAACCATCTCGAGAGAACTCAGAGATCCACA 1570
2119 GAAAAGACCTTTAGCGAGAGCTCCATCTTAAGTCAATCAAGAAATTCACA 2170

RESULT 9
US-10-104-047-693
Sequence 693, Application US/10104047
Patent No. 6943241
GENERAL INFORMATION:
APPLICANT: HELIX RESEARCH INSTITUTE
TITLE OF INVENTION: No. 6943241el full length cdna
FILE REFERENCE: H1-A0105
CURRENT APPLICATION NUMBER: US/10/104,047
PRIOR FILING DATE: 2002-03-25
NUMBER OF SEQ ID NOS: 4096
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 693
LENGTH: 2241
TYPE: DNA
ORGANISM: Homo sapiens
US-10-104-047-693

Query Match 23.6%; Score 474.8; DB 3; Length 2241;
Best Local Similarity 65.9%; Pred. No. 5.4e-140;
Matches 705; Conservative 0; Mismatches 362; Indels 3; Gaps 1;

422 TTGGCCGTACTGATATGTGAAGTACGCTTGAAGAACCAAGCGCTCCAGAGGAGGT 481
1169 TTGGAGATTTTGGAAATTAAGACCACTTGCAAAACAATGGATATCCAAATGGGC 1228
482 ATGTCAATGATGATCAATTAATGTCAAAAAGCTGTCTATAGAGAGGACCCCTC 541
1229 ATTTATGTAAGAAATATTCATCTCGAATATACGCCCCATTTATTCACAGACATTC 1288
542 CTAGAACATCAAG 598
1289 TTAATCTCATCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1348
599 AGGCTTATGAGTGTGCTTCACTTATGATCAATCAAGAAATCCATCTGTGTGAGAAAC 658
1349 AGGCTTATGAGTGTGCTTCACTTATGATCAATCAAGAAATCCATCTGTGTGAGAAAC 1408
659 CTATGATGTAAAGAGTAAAGAGGCTTCCGTTGGGCAATCACTTACTCAATC 718
1409 CTATGATGTAAAGAGTGTGAGAAATCTTTATGTTGTGTTCACATGTTACTCGGCATC 1468
719 AAAAAATTCATCTGGGAGAGAGCCCTTACAGATGTAAAGCTGTGGAGAGGCTTTCCAT 778
1469 TGAATTTATCTGTGCGAGAAACCTTGAATGTAAAGAGTGTGAAAGGCTTCAAGTT 1528
779 GGGGCTCAAGCTCGTTATTCATTAAGAGATTCATCTGTGTGAAAACCTTATGATTA 838
1529 GTAGCTCATACCTTTTCTCAACATCAGAGAAATCCTATGCGGTAAAGAAACCTTATGATTA 1588
839 AAGACTGTGAAAAGGCTTTCGCGGTGTGATGAGTCACTCAGCAAGAGATTCACAG 898
1589 AGGAATGTGGAAGAGGCTTTATGTTATGTCTCAATCTTATGACATCAAGCAATTCACA 1648
899 CTGGGAGAAAGAGCTACAGATGCAAGAGCTGTGTGAGAAAGCTTTAGCCGTGTATTAAC 958
1649 CTGGTAAAAACCTTATGATGTAAAGTGTGTGAGAAAGCTTTATGATTAAGAGCTCAAC 1708
959 TTAATGAGCAAGAGATTCATGATGTGTGAGAGAGGCTTACAGATGTAAAGAGCTGTGGA 1018
1709 TTTTTCAGATGACGAATTCATCAGGTGTGAGAGAAACCTTATGATTAAGAGATGTGGA 1768
1019 AGGCTTTATTTGTGTTCAGGCTTATCAGCATTAAGAAATTCACAGAGGTGAGAAAC 1078


```
Db 1769 AAGCCTTTACCCAGAGCTCAAAAGCTTGTTCACATCAGAGAAATTCATCTGGTGAAGAAC 1828
Qy 1078 CCTATGATGTCAAGATGTGGAGAGCCTTTACTGAGTCAATTAACCTTACTCAGATC 1138
Db 1829 CCTATGATGTCAAGATGTGGAGAGCCTTTACTGAGTCAATTAACCTTACTCAGATC 1888
Qy 1139 AGAAGATTCACACCGGTGAGAAAGCCTCAGAAATGTAGAGATGTGGAGAGCCTTTCGCT 1198
Db 1889 AGAAGATTCACACCGGTGAGAAAGCCTCAGAAATGTAGAGATGTGGAGAGCCTTTCGCT 1948
Qy 1199 GGGGTTGAGCCTCGTTTAAAGCAGAGAGATACATCGGGCGAGAGCCGTACAAGTGA 1258
Db 1949 AAGAGTCACAGCTCTGCTCAACATCAGAGAAATTCACCTGCTGAGAGAACCTTTGAGATGTC 2008
Qy 1259 CAGATGTGGAGAGCCTTCAATTTGCTATCACTCACTCAGCAGAGAAATTCACA 1318
Db 2009 TTGAATGTGGAGAGCCTTACTCAGAACTCACAATTTTCCAGATCAGAGAAATTCATA 2068
Qy 1319 CAGGGGAAACCCGTTAATGTAAAGAGATGTGGAGAGCCTTCAATTAATGTAGATCAGAGC 1378
Db 2069 CAGATGAAACCCATATGATGTAAATGTAGAAAGCCTTTAAATTAATGCTCAAAAC 2128
Qy 1379 TCGTGAACATGAGAGAAATTCATACCGGGGTGAAACCTATGGGTGTAGAGATGTGGGA 1438
Db 2129 TTAATCGATCTGAGAAATTCACACTGCTGAGAAAGCCTTAATGCTTAAAGATGTGGGA 2188
Qy 1439 AAGAGCTTATGTCACGGCCATCAGCTTACACAAATCAGAGAAACGACAGT 1488
Db 2189 AAGCTTTATGATGTGGCTCGAGATTCATTCATTCATCAGAGAAATTCATACT 2238
```

RESULT 10

```
US-10-104-047-1464
; Sequence 1464, Application US/10104047
; Patent No. 6943241
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: No. 6943241el full length cdna
; FILE REFERENCE: H1-A0105
; CURRENT APPLICATION NUMBER: US/10/104,047
; PRIOR FILING DATE: 2002-03-25
; PRIOR APPLICATION NUMBER:
; NUMBER OF SEQ ID NOS: 4096
; SOFTWARE: Patent Ver. 2.1
; SEQ ID NO 1464
; LENGTH: 1914
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-104-047-1464
```

```
Query Match 22.9%; Score 461.6; DB 3; Length 1914;
Best Local Similarity 66.5%; Pred. No. 7.8e-136;
Matches 662; Conservative 0; Mismatches 334; Indels 0; Gaps 0;
```

```
Qy 575 CCTTGAATGTAAAGACCTGTGGAGAGCCTTTAGTGTGCTATCAACTTAAGTCAATC 634
Db 793 CCTTGAATGTAAAGACCTGTGGAGAGCCTTTAGTGTGCTATCAACTTAAGTCAATC 852
Qy 635 AGAAATCATCTGCTGAGAAAGCCTTATGAAATGTAAAGATGAAGAGCCTTCCGT 694
Db 853 AGCAATTCATCTGAGAGAAAGCCTTATGAAATGTAAAGATGAAGAGCCTTCCAGTA 912
Qy 695 GGGGCAATCAGCTTACTCAATCAATCAAAATTCATCTGGGAGAGAGCCTTACGATGA 754
Db 913 GGAAGGAAATCTTATTAACATCAGAAATTCATCTGGGAGAGAGCCTTAAAGTGA 972
Qy 755 AAGACTGTGGAGAGCCTTTTCATGAGGCTCAAGCTCTGTTATTCATTAAGAGATTCATA 814
Db 973 ATGAATGTGGAGAGCCTTTTCATTCAGATGTCAAACTTTATTTAGACCCACAGAAATTCATA 1032
Qy 815 CTGTGAGAAACCCATATGATGTAAAGACTGTGAGAAAGCCTTTGGCGTGTGATGAGC 874
```

```
Db 1033 CTGGGAGAGAACTTATGATGTAAAGATGTGTGAGAAAGCCTTACAGTCAAGAAATCAATC 1092
Qy 875 TCACTCAGACAGAGATTCACACTGGGAGAGAAAGCTACAAATGCAAAAGCTGTGGGA 934
Db 1093 TCAATGAAACATGAGGAAATTCACACTGAGAGAGAAAGCCTTATGAAATGTAAAGATGTGGGA 1152
Qy 935 AAGCCTTTAGCCGTGTGTATTAATTAATTCAGACAGAGAAATTCATATGTGGGAGAGAAC 994
Db 1153 AATCTTCAGCCAGAGAGAAATCTTATGAGACAGAGAAATTCATATCTGGGAGAGAAC 1212
Qy 995 CTTACAGATGTAAAGACTGTGGAGAGCCTTTATTTGTGTTCAAGCCTTATTCAGACATA 1054
Db 1213 CTTATGATGTATATGATGTGTAGAGCCTTTTCTCGAATGTCACTGTGTAGCTACATTA 1272
Qy 1055 AAGAAATTCACACAGGTGAGAAACCTTATGAAATGTCAAGATGTGGAGAGCCTTACTC 1114
Db 1273 TGAGAAATTCACACAGGGAGAGAAACCTTATTAATGTAAATGTGAGAGAAAGCCTTCTC 1332
Qy 1115 GAGTCAATTCCTTACTCAGATCAGAAATTCACACCGGTGAGAGAGCCTTCAAGATGA 1174
Db 1333 AATGCTCAGTATTTATTTATCAATATGAGAAATGACACCTGTGAGAGAAACCTTATGATGA 1392
Qy 1175 AAGAGTGTGGAGAGCCTTTGCTGGGTGAGAGCCTGTTAAGCAGAGAGATACATA 1234
Db 1393 GTGAATGTGGAGAGCCTTCTCTCAGAGATTCATCCCTTAACCTATATGAGAAATCATA 1452
Qy 1235 CGGGGAGAGAGCCTTACAGAGTGCAGAGATGTGGAGAGCCTTCAATGTGTGCTATCACC 1294
Db 1453 CAGCTGAGAGAAACCTATGAAATGTAAAGATGTGAGAAAGCCTTACAGAGAGAAAGAAATC 1512
Qy 1295 TCACTCAGACAGAGAGATTCACACAGGCGAGAAACCCGTATTAATGTAAAGATGTGGGA 1354
Db 1513 TCATTACATCAGAGAAATTCACACTGAGAGAGAAACCTTATGAAATGAGTGTGAGGA 1572
Qy 1355 AAGCTTTCAATTTATGATGTGAGCTGTGAAACATGAGAGAAATTCATACCGGGGTGAAC 1414
Db 1573 AAGCTTTATTCAGATGTCAACCTTCAATTCGACACAGAGAAATTCATACGGGTGAGAAAC 1632
Qy 1415 CCTATGAGTGTACAGAAATGTGGAGAGCCTTTAGTCAAGCCATCAGCTTACACAAATC 1474
Db 1633 CCTATGAGTGTACAGAAATGTGGAGAGCCTTTAGTCAAGAAATCAAACTCAGTGAACATG 1692
Qy 1475 AGAAAGCGACAGTGGGCGAAATCTTACGAATGTAAAGAGTGGGAGAGCAGTGAAC 1534
Db 1693 AGAAATTCATCTGAGAGAGAAACCTTATCATTTGTAATCAATGTGGAGAGCCTTCAATC 1752
Qy 1535 ACCTAAACATCTCGAGAGACATCAGAGATTCACA 1570
Db 1753 AGAGCAAAATCTTGTAGCATGAGAAATTCATA 1788
```

RESULT 11

```
US-09-949-016-4204
; Sequence 4204, Application US/09949016
; Patent No. 6812359
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: C1001107
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4204
; LENGTH: 3798
```

TYPE: DNA
ORGANISM: Human
US-09-949-016-4204

Query Match 22.9%; Score 460.6; DB 3; Length 3798;
Best Local Similarity 64.9%; Pred. No. 2,5e-135;
Matches 682; Conservative 0; Mismatches 369; Indels 0; Gaps 0;

```

QY 535 ACCCTCTAGAAACATCATGAGACATCATTAAGAAATTCCTTTAATGTAAGACTGT 594
DB 1816 ACCCTTAATGACATTAAGAGAGACACATCGAGAGAAACCTTACAAATGTGAAGATGT 1875
QY 595 GGGAGGCGCTTTAGTGTGCTATCACTTAAGTCAATCAAGAAATCCATCTGTGTAG 654
DB 1876 GGGAAAGCTTTTATGACATTTCTTCAACCTTGTCAAAATGAAGAAATTCATCTGAGAG 1935
QY 655 AAACCTTAAGATGTAAGAAATGTAAGAGGCTTCCTGTTGGGCAATCAGTACTTCA 714
DB 1936 AAACCTTAAGATGTAAGAAATGTAAGAGGCTTCCTGTTGGGCAATCAGTACTTCA 1995
QY 715 CATCAAAAAATTCATCTAGTGGGAGAAAGCCCTAAGAAATGTAAGAGTGTGGAGGCTTT 774
DB 1996 CATTAAGAAATTCATCTAGTGGAGAAAGCCCTAAGAAATGTAAGAGTGTGGAGGCTTT 2055
QY 775 CGATGGGCTCAAGCCTCTTATTCATTAAGAGATTCATCTGTGAAAAACCTTATGAA 834
DB 2056 AGCAATTTCTCAACCTTGTCTATCATTAAGATTCATCTGTGAAAAACCTTATGAA 2115
QY 835 TGTAAAGATCTGTGAGAAAGCCTTTCGGCTGTGATGAGTCACTAGCAACGAGATTC 894
DB 2116 TGTAAAGATCTGTGAGAAAGCCTTTCGGCTGTGATGAGTCACTAGCAACGAGATTC 2175
QY 895 CACACTGGGAGAGAACTACGATCGAATCAAGCTGTGGAGAGCCTTTAGCCGTGTAT 954
DB 2176 CAGCTGTGAGAGAACTCTTCAATGTGAAGATGTGGCAAGCCTTTATTCATCTTCA 2235
QY 955 AAACCTTATTCAGACAAGAAATTCATAGTGGGAGAGCCTTTCAGATGTAAAGACTGT 1014
DB 2236 AAACCTTATTCATTAAGATTTTATTCATCTGAGAGAAAGCCTTTCAGATGTAAAGATGT 2295
QY 1015 GGGAGGCTTTTATTTGTGTTCAGGCTTCATCAGATTAAGAAATTCACAGGTGAG 1074
DB 2296 GGGAAAGCATTTTATCTGTCTCAAGCCTTTCATTAAGAAATTCATCTAGAGAG 2355
QY 1075 AAACCTTATGATGTCAAGATGTGGAGAGCCTTTTACGATGCAATTCCTTACG 1134
DB 2356 AAACCTTCAAAATGTAAGAAATGTGGCAAGCATTTTATGTCTTCAACCTTACAG 2415
QY 1135 CATCAGAAATCCACACCGGTGAGAAAGCCTCAGATGTAAAGAGTGTGGAGGCTTT 1194
DB 2416 CATTAAGAGATCACACTGGAGAGAGCCTTACAAATGTGAAGATGTGGCAAGCTTT 2475
QY 1195 CGGTGGGCTTCAAGCCTCGTTAAGCAAGAGATTCATACCGGCGAGAGCCGTCAAG 1254
DB 2476 AGCGTTTCTCAACCTTACTTAAGCATTAAGCAATTCATCTGAGAGAAAGCCTTACAA 2535
QY 1255 TGCACAAATGTGGAGAGCCTTCAATTTGGCTTATCACTCACTCAGCAGAGATTC 1314
DB 2536 TGTAAAGAAATGTGGCAAGCTTTTAAAGCATCTCTCAGCCTTGTGTAACATTAATA 2595
QY 1315 CACACAGGCGAAACCCGTATTAATGTAAAGAGTGTGGAGAGCTTTCAATTTATGATCG 1374
DB 2596 CATGTGTGAGAGAACTCTCAATATGTGAAGAAATGTGGCAAGCTTTTATTCATCTTCA 2655
QY 1375 AGCTCTGTAAATGATGAGAAATTCATACCGGGGTGAACCTTATGGGTGTACAGATGT 1434
DB 2656 AATCTTACGACATTAAGATTAATTCATTAAGAGAAACCTTCAAGAGTGAAGATGT 2715
QY 1435 GGGAGAGCTTTTATGTCAGGCGCATCAGCTTACACAAATCATGAGAAAGCAGAGGCGC 1494
DB 2716 GAGAAAGCATTTTATGTGTCTTCAACCTTATCTGAACATTAAGAAATTCATCAAGAG 2775
QY 1495 AAATCTTACGAATGTAAAGAGTGTGGAGAGCATGTAAACCATCTTCCGAGAA 1554

```

```

DB 2776 AAACCTTAAGATGTAAGAAATGTGGCAAGATTTAGCACCTTACACCTTACTACA 2835
QY 1555 CATCAGAGATCCACACAGCTTGAAGAGCCT 1585
DB 2836 CATTAAGAGATGCACACTGGAGAGAAACCTT 2866

```

RESULT 12

```

US-09-949-016-12227
; Sequence 12227, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 12227
; LENGTH: 156942
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1)..(156942)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-12227

```

Query Match 22.9%; Score 460.6; DB 3; Length 156942;
Best Local Similarity 64.9%; Pred. No. 2.7e-134;
Matches 682; Conservative 0; Mismatches 369; Indels 0; Gaps 0;

```

QY 535 ACCCTCTAGAAACATCATGAGACATCATTAAGAAATTCCTTTAATGTAAGACTGT 594
DB 152967 ACCCTTAATGACATTAAGAGAGACACATCGAGAGAAACCTTACAAATGTGAAGATGT 153026
QY 595 GGGAGGCGCTTTAGTGTGCTATCACTTAAGTCAATCAAGAAATCCATCTGTGTAG 654
DB 153027 GGGAAAGCATTTTATCTGTCTCAAGCCTTTCATTAAGAAATTCATCTAGAGAG 153086
QY 655 AAACCTTATGATGTAAAGATGTAAAGAGCCTTTCGTTGGGCAATCAGCTTACTCA 714
DB 153087 AAACCTTACAAATGTGAAGATGTGGCAAGCTTTTACCGTTCCTTCAACCTTGTCAAA 153146
QY 715 CATCAAAAAATTCATCTAGTGGGAGAGCCCTTACGATGTAAAGAGTGTGGAGGCTTT 774
DB 153147 CATTAAGAAATTCATCTGAGAGAAAGCCTTCAAAATGTGAAGAAATGTGGCAAGCTTT 153206
QY 775 CGATGGGCTCAAGCCTCGTTATTCATTAAGAGATTCATCTGTGAAAAACCTTATGAA 834
DB 153207 AGCAATTTCTCAACCTTGTCTATTCATTAAGATTCATCTGAAGAAACCTTACAA 153266
QY 835 TGTAAAGATGTGAGAAAGCCTTTCGGCGTGTGATGAGTCACTAGCACAGATTC 894
DB 153267 TGTAAAGATGTGAGAAACCTTTTAAAGCATCTCAACCTTACTTAACATTAATAATA 153326
QY 895 CACACTGGGAGAGAAAGCATTCAGAAATGCAAGAGCTGTGGAGAGCCTTATGGCTGTGTAT 954
DB 153327 CATGTGTGAGAGAACTCTCAATATGTGAAGAAATGTGGCAAGCTTTTATTCATCTTCA 153386
QY 955 AAACCTTATTCAGACAAGAAATTCATAGTGGGAGAGAGCCTTACAGATGTAAAGACTGT 1014
DB 153387 AATCTTATCTTACATTAAGATTTTATTCATCTGGAGAGAAACCTTTCAGAGTGTGAAGATGT 153446

```


RESULT 14
US-10-104-047-271
; Sequence 271, Application US/10104047
; Patent No. 6943241
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: No. 6943241el full length cDNA
; FILE REFERENCE: H1-A0105
; CURRENT APPLICATION NUMBER: US/10/104,047
; CURRENT FILING DATE: 2002-03-25
; PRIOR APPLICATION NUMBER:
; PRIOR FILING DATE:
; NUMBER OF SEQ ID NOS: 4096
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 271
; LENGTH: 2669
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-104-047-271

Query Match 22.7%; Score 456.6; DB 3; Length 2669;
Best Local Similarity 65.4%; Pred. No. 3.8e-134;
Matches 669; Conservative 0; Mismatches 354; Indels 0; Gaps 0;

QY 568 GAGAAATTCCTTGAATGTAAGACTGTGGAGAGCCCTTAGTCGTGCTATCACTTAGT 627
DB 1159 GAGAAACCTTAGAGAGTGTATGACTGTGGAGAGCCCTTAGTCGTGTAATGCAAAATTAATT 1218
QY 628 CAACATCAGAAATTCATCTAGTGTGAGAAACCTTAGTAATGTAAGAAATTAAGAGCC 687
DB 1219 CAACATCAGAAATTCATCTAGTGTGAGAAACCTTAGTAATGTAAGAAATTAAGAGCC 1278
QY 688 TTTCCGTTGGGGCAATCAGCTTACTCAACATCAAAATTAATCTGSGGAGAGCCCTAC 747
DB 1279 TTCAAGTGAAGTCCAGCTTACAGCTTACAGCATCAACAGAGAAAGCCCTAT 1338
QY 748 GAATGTAAAGACTGTGGAGAGCCCTTTCGATGGGCTCAAGCCTGTAATCAAGAGG 807
DB 1339 CAGTGTAAAGAGTGTGAGAAAGCCCTCAATTAATTAATCAAAATCAATTCAGATCAAG 1398
QY 808 ATTCACTACTGTGTGAGAAACCTTATGATGTAAAGACTGTGGAGAGCCCTTTCGCTGT 867
DB 1399 ATCCACACAGTGTGAGAAACCTTATGATGTCACTGAATGTGAGAAAGCCCTTCACTGTC 1458
QY 868 GATGAGCTCAGCTGAGACCGAGATTCACACTGGGGAGAGAACTACGAATGCAAGAG 927
DB 1459 GGGAACTTAATTCAGACCGAGAAATTCACACAGCGAGAAACCTTAGTGTATGAA 1518
QY 928 TGTGGAGAGACTTTAGCCGTGTGTATTAATTCAGACAGAGAAATTCAGTGGG 987
DB 1519 TGGGGAGAGCCCTTACAGATGTAACTCCCAATTCGGAGAGATCGAGAAATTCACACTGG 1578
QY 988 GAGAGCCCTTACAGTGTAAAGACTGTGGAGAGCCCTTTATTTGTGTTCAAGCTCAT 1047
DB 1579 GAGAGCCCTTATAGTGTATGTAGTGTGAGAAAGCCCTTTCAGCTTATGAGAAACTA 1638
QY 1048 CAGCATTAAGATTCACACAGGTGAGAAACCTTATGATGTCAAGATGTGAGAGAGCC 1107
DB 1639 CGGCACTACAGATTCACACTGGAGAGAACTTTTGAATGTATAGTGTGGAGATGC 1698
QY 1108 TTTACTCGAGTCAATTACTTACAGCATCAGAGATTCACACCGGTGAGAGCCCTCAC 1167
DB 1699 TTTACTCTTAAGAAAGAACTTATGATGTATCAACGATTCATCTGAGAGAAAGCCCTAT 1758
QY 1168 GAATGTAAAGAGTGTGGAGAGCCCTTTCGCTGGGTTTCGAGCTCGTTAAGCAGAGAG 1227
DB 1759 CATGTAAAGAAAGTGTGGAGAGCCCTTTCAGTATCAATCCAACTTAACAGCATCAGAG 1818
QY 1228 ATACATACGGGAGAGAGCCGTACAGTACAGAGATGTGAGAGAGCCCTCAATTGAGG 1287
DB 1819 ATTCAATCTGGGAGAGAACTTTCAATGTATGTGAAATGTAGAGAGATTCAGCTGTAGT 1878

QY 1288 TATCACTTCACTCAGCAGAGAGAAATTCACACAGCGAGAAACCCGTATTAATGTAGAG 1347
DB 1879 TCTAATTAATGTGTACACAGAGAAATTCATACAGAGAGAAACCCCTTTCAGTGTAGAG 1938
QY 1348 TGTGGAGAGCCCTTCAATTAATGTATGATCGAGCCCTGTGAAACATGAGAGATTCATCCGG 1407
DB 1939 TGTGAGAAAGCCCTTCAATGTATGTCCCATTTAATTCGGCATCAGAGAGCCCATCTGG 1998
QY 1408 GTGAAACCTTATGGTGTACAGAAATGTGGAGAGAGCTTATGACGGCCATCAGCTTACA 1467
DB 1999 GAGAAACCTTCAAGTGTGTGAAATGTGGCAAGGCTTCAGCTTATGTTCTGACTACATT 2058
QY 1468 CAACATCAGAAACGACAGTGGGCGGAAATCTTACAAATGTAAGAGTGTGGAGAGCA 1527
DB 2059 ATACACAGACAGTCCACACTTGGAGAAACCTTATGTGTGTGTGTGTGGAGAGCA 2118
QY 1528 TGTAAACCTTAACCATCTCCGAGAACATCAGAGATTCACAAAGTTGAAGAGCCCTT 1587
DB 2119 TTCAAGTTTACCTTCCAGCTCAGTCACTCAGATGTCTCATGTGAGAGAAATCTTAA 2178
QY 1588 TGA 1590
DB 2179 TAA 2181

RESULT 15
US-09-949-016-485
; Sequence 485, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 485
; LENGTH: 3839
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-485

Query Match 22.6%; Score 455.2; DB 3; Length 3839;
Best Local Similarity 65.0%; Pred. No. 1.3e-133;
Matches 673; Conservative 0; Mismatches 363; Indels 0; Gaps 0;

QY 535 ACCCTTCCTAAGACACATCAGACATCATTAAGAGAAATTCCTTGAATGTAGAGCTGT 594
DB 1688 ACCCTTAATTAACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1747
QY 595 GGGAGAGCCCTTATAGTGTGCTATCAACTTAAGCAATCAAGAAATTCATCTGGTAG 654
DB 1748 GGGAGAGCCCTTATAGCAATTCACCTTACTACACATTAATTAATTAATTAATTAAT 1807
QY 655 AAACCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 714
DB 1808 AAACCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1867
QY 715 CATCAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 774
DB 1868 CATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1927
QY 775 CGATGGGCTCAAGCCCTGTATTAATTAAGAGATTCATCTGTTGAAAAACCTTGA 834
DB 1928 CTATGCTTCAACCTTAAGAGACATTAAGAGATTAACCTGAGAGAAACCTTACAA 1987

THIS PAGE BLANK (USPTO)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceeleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 6, 2006, 16:01:11 ; Search time 1223 Seconds
(without alignments)
10980.667 Million cell updates/sec

Title: US-10-111-257-3

Perfect score: 2015

Sequence: 1 cgctctgcacacgggtgacgc.....gtgaagttcctaacttaa 2015

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4996997 seqs, 3332346308 residues

Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N_Geneseq_21:*

1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002as:*
7: geneseqn2002bs:*
8: geneseqn2003as:*
9: geneseqn2003bs:*
10: geneseqn2003cs:*
11: geneseqn2003ds:*
12: geneseqn2004as:*
13: geneseqn2004bs:*
14: geneseqn2005s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	2015	100.0	2015	5	AAf88159	AAf88159 Human thy
2	1986.8	98.6	2098	4	AA503044	AA503044 Human dia
3	1967.4	97.6	2173	5	ABa82974	ABa82974 Human tra
4	1958	97.2	2169	4	AAK94562	AAK94562 Human ful
5	1946.2	96.6	3293	4	AA625944	AA625944 Human nov
6	1946.2	96.6	3293	8	ABX73285	ABX73285 Human nov
7	1946.2	96.6	3293	8	ABX73285	ABX73285 Human nov
8	1768	87.7	2010	6	ABH59917	ABH59917 Novel hum
9	1420.2	70.5	1453	4	AAH99589	AAH99589 Human pro
10	1397.8	69.4	1429	11	AD131363	AD131363 Human cdn
11	1397.8	69.4	1429	13	AD883430	AD883430 Human lym
12	734.6	36.5	4483	10	ADP14367	ADP14367 Human end
13	703	34.9	2143	6	ABa92946	ABa92946 Human car
14	703	34.9	2143	8	ACA62107	ACA62107 CDNA enco
15	620.6	30.8	2832	10	ADC30401	ADC30401 Human nov
16	620.6	30.8	3547	6	ABH84143	ABH84143 Human zin
17	618	30.7	1945	10	ADA53552	ADA53552 Human cod
18	615.4	30.5	2235	11	ADM03674	ADM03674 Human cdn
19	603.6	30.0	1900	5	AA589495	AA589495 DNA enco

20	602	29.9	2617	4	AAK53340	AAK53340 Human pol
21	600.4	29.8	2650	4	AAK52356	AAK52356 Human pol
22	593.6	29.5	774	8	AA526388	AA526388 Human cdn
23	593.6	29.5	774	8	ABX73729	ABX73729 Human nov
24	584.4	29.0	592	12	ACH79967	ACH79967 Human gen
25	575.6	28.6	1521	10	AA155798	AA155798 CDNA enco
26	573.8	28.5	1749	13	AD086903	AD086903 Human tum
27	548.8	27.3	2223	4	AB196293	AB196293 New zinc
28	548.6	27.2	2276	4	AA160063	AA160063 Human pol
29	546	27.1	1335	12	ADN99068	ADN99068 Novel hum
30	546	27.1	1335	12	ADN99068	ADN99068 Novel hum
31	544.4	27.0	1970	13	ADP54658	ADP54658 Human PRO
32	538.6	26.7	1813	5	ABA83028	ABA83028 Human tra
33	538	26.7	2334	10	ADA53517	ADA53517 Human cod
34	538	26.7	4683	5	AA568826	AA568826 DNA enco
35	520.6	25.8	2274	5	AA158277	AA158277 Human pol
36	520.6	25.8	2274	5	ADQ98484	ADQ98484 DNA enco
37	520.6	25.8	2274	5	ADQ98484	ADQ98484 DNA enco
38	512	25.4	5718	11	ABE66032	ABE66032 Novel hum
39	510.4	25.3	5718	13	ACN40550	ACN40550 Tumour-as
40	510.2	25.3	526	6	ABL63275	ABL63275 Breast ca
41	510	25.3	1874	8	ABX71104	ABX71104 Novel hum
42	509.2	25.3	1188	8	ABX34452	ABX34452 Human md
43	508.8	25.3	2332	10	ADC30875	ADC30875 Human nov
44	507.6	25.2	3358	9	ACA98941	ACA98941 CDNA enco
45	506.8	25.2	2624	4	AAK94790	AAK94790 Human ful

ALIGNMENTS

RESULT 1	AAf88159	standard; CDNA; 2015 BP.
ID	AAf88159	
XX	AAf88159;	
AC		
DT	17-JUL-2001	(first entry)
XX		
DE	Human thyroid malfunction-associated protein RITA encoding CDNA.	
XX		
KW	KRAB domain; hyperplasia; thyroid; tumor; zinc finger motif; cyclostatic; antithyroid; gene therapy; chromosome 19; 19q13; ss.	
XX		
OS	Homo sapiens.	
XX		
FH	Key	Location/Qualifiers
FT	CDS	187..1578
FT		/*tag= a
FT		/product= "RITA"
XX		
PN	WO200127265-A1.	
PD	19-APR-2001.	
XX		
PF	11-OCT-2000; 2000WO-DE003600.	
XX		
PR	12-OCT-1999; 99DE-01049179.	
XX		
PA	(UWBR-) UNIV BREMEN.	
XX		
PI	Bulleriek J, Rippe V, Meiboom M, Belge G;	
XX		
DR	WPI, 2001-290723/30.	
XX		
DR	P-PEDB; AAB86115.	
XX		
PT	New nucleic acid useful for the diagnosis and treatment of thyroid disorders, e.g. tumors.	
XX		
PS	Claim 2; Page 52-54; 59pp; German.	
XX		
CC	This invention describes a novel nucleic acid (N1) encoding a polypeptide which comprises a KRAB-domain and/or at least one zinc finger motif. The	

CC products of the invention have cytostatic and antithyroid activity and
CC can be used in gene therapy. Nucleic acids, polypeptides, and antibodies
CC of the invention may be used in the diagnosis and/or the therapy of the
CC malfunction of the thyroid and/or hyperlasias of the thyroid and/or
CC thyroid tumors. They may also be used in the production of medicaments.
CC (N1) can also be used to diagnose thyroid tumors which are located on
CC chromosome 19 at band 19q13. This sequence encodes the thyroid
CC malfunction-associated protein, R19A which is described in the method of
CC the invention
XX

Sequence 2015 BP; 601 A; 444 C; 503 G; 467 T; 0 U; 0 Other;

Query Match 100.0%; Score 2015; DB 5; Length 2015;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 2015; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 CGCTTTGACCGGTGATGCGCAACCGCTGTGTCTCCGCCAGTCCGCCAGGCCAGCATCTT 60
DB 1 CGCTTCTGACCGGTGACCGCAACCGCTGTGTCTCCGCCAGTCCGCCAGGCCAGCATCTT 60
QY 61 TCAGAAAAAGCATCCCGAGAGAGAGAGAGATCGTTAAACATCTTAGCTCAGCTTAGC 120
DB 61 TCAGAAAAAGCATCCCGAGAGAGAGAGAGATCGTTAAACATCTTAGCTCAGCTTAGC 120
QY 121 CTCTCGAAATTTGTCTTTCTTCACTGAGAAACCCCGAGAGAGATGATCAGTTCTTCACTTCT 180
DB 121 CTCTCGAAATTTGTCTTTCTTCACTGAGAAACCCCGAGAGAGATGATCAGTTCTTCACTTCT 180
QY 181 AAAACATATGGCCAGAGGTTTGTGACGTTCCGCGAGATGACCTATACCTTTTTCAGAG 240
DB 181 AAAACATATGGCCAGAGGTTTGTGACGTTCCGCGAGATGACCTATACCTTTTTCAGAG 240
QY 241 GAGTGGGCGCTGTCTGAACTCTGCTCAGAGGAGACCTGATCTGAGAGCGTGTGAGAGAC 300
DB 241 GAGTGGGCGCTGTCTGAACTCTGCTCAGAGGAGACCTGATCTGAGAGCGTGTGAGAGAC 300
QY 301 TACAGTAACTTGTGTCTCACTGATTTTGAATGAGATGAGAAATTAAGATTACTTACA 360
DB 301 TACAGTAACTTGTGTCTCACTGATTTTGAATGAGATGAGAAATTAAGATTACTTACA 360
QY 361 GAAAAAAACATTTATGAAATTAAGGCTTCCAAAAGAAATTCAGATGAGAAAGTAATCC 420
DB 361 GAAAAAAACATTTATGAAATTAAGGCTTCCAAAAGAAATTCAGATGAGAAAGTAATCC 420
QY 421 CTGGGCGTAACTGATGATGATGAGGATGAGCTTGAAGAACACAGGCGTCCAGAGGAG 480
DB 421 CTGGGCGTAACTGATGATGATGAGGATGAGCTTGAAGAACACAGGCGTCCAGAGGAG 480
QY 481 TATGTCATTCAGATGATCATCAATTAATGCAAAAGGCTGTCTAATAAGAGGACCCCT 540
DB 481 TATGTCATTCAGATGATCATCAATTAATGCAAAAGGCTGTCTAATAAGAGGACCCCT 540
QY 541 CCTAGAACATTCAGAGACATCATTAAGAGAAATTTCTTGAATGTAAGAGCTGTGGAG 600
DB 541 CCTAGAACATTCAGAGACATCATTAAGAGAAATTTCTTGAATGTAAGAGCTGTGGAG 600
QY 601 GCCTTAGTGTGGCTATCAACTTAAGTCAATCAGAAATTCATCTGTGTGAGAAACCT 660
DB 601 GCCTTAGTGTGGCTATCAACTTAAGTCAATCAGAAATTCATCTGTGTGAGAAACCT 660
QY 661 TATGAATGTAAAGATGTAAGAGGCTTCGTTGGGCAATCAGCTTACTCAACATCA 720
DB 661 TATGAATGTAAAGATGTAAGAGGCTTCGTTGGGCAATCAGCTTACTCAACATCA 720
QY 721 AAAATTCATCTGTGGAGAGGCTTACGAAATGTAAGAGCTGTGGAGGCTTTTCATGG 780
DB 721 AAAATTCATCTGTGGAGAGGCTTACGAAATGTAAGAGCTGTGGAGGCTTTTCATGG 780
QY 781 GGTCTCAGGCTGTGTTATTCATTAAGAGATTCATCTGTGAAAGAAACCTTATGATGAA 840
DB 781 GGTCTCAGGCTGTGTTATTCATTAAGAGATTCATCTGTGAAAGAAACCTTATGATGAA 840
QY 841 GACTGTGAAAGGCTTTCCGCGTGTGATGAGCTCATCAGCACAGAGATTCACACT 900
```

```
DB 841 GACTGTGAAAGGCTTTCCGCGTGTGATGAGCTCATCAGCACAGAGATTCACACT 900
QY 901 GGGGAGAAAGCTACGAATGCAAGACGTGGGAGACCTTAGCCGTGTGTAATAACTT 960
DB 901 GGGGAGAAAGCTACGAATGCAAGACGTGGGAGACCTTAGCCGTGTGTAATAACTT 960
QY 961 ATTCAGACAAGAGAAATTCATAGTGGGAGAGCCTTACGAGTGTAAAGACTGTGGAG 1020
DB 961 ATTCAGACAAGAGAAATTCATAGTGGGAGAGCCTTACGAGTGTAAAGACTGTGGAG 1020
QY 1021 GCTTTTATTTTGTGTTCAAGCTCATTCAGCTTAAAGAAATTCACAGGTGAGAAACC 1080
DB 1021 GCTTTTATTTTGTGTTCAAGCTCATTCAGCTTAAAGAAATTCACAGGTGAGAAACC 1080
QY 1081 TATGAATGTCAAGATGTGGAGAGGCTTTACTCGAGTCAATTAATCTTACTCAGATCAG 1140
DB 1081 TATGAATGTCAAGATGTGGAGAGGCTTTACTCGAGTCAATTAATCTTACTCAGATCAG 1140
QY 1141 AAGATCCACACCGGTGAGAGGCTTCAGAAATGTAAGAGAGTGTGGAGAGGCTTTGCTGG 1200
DB 1141 AAGATCCACACCGGTGAGAGGCTTCAGAAATGTAAGAGAGTGTGGAGAGGCTTTGCTGG 1200
QY 1201 GGTTCAGGCTGTGTTAAGCAGAGAGATACATACGAGGAGAGAGCCGTTAAGAGCCACA 1260
DB 1201 GGTTCAGGCTGTGTTAAGCAGAGAGATACATACGAGGAGAGAGCCGTTAAGAGCCACA 1260
QY 1261 GAAATGGGAGAGGCTTCATATGTGGCTATCACTCACTCAGACAGAGAAATCCACACA 1320
DB 1261 GAAATGGGAGAGGCTTCATATGTGGCTATCACTCACTCAGACAGAGAAATCCACACA 1320
QY 1321 GCGGAAACCCCGTATTAATGTAAGAGATGTGGAGAGGCTTTCAATTAATGATCGAGCTC 1380
DB 1321 GCGGAAACCCCGTATTAATGTAAGAGATGTGGAGAGGCTTTCAATTAATGATCGAGCTC 1380
QY 1381 GTGAAACATGAGAAATTCATACCGGGGTGAAACCTTATGGGTATCAGAAATGTGGAG 1440
DB 1381 GTGAAACATGAGAAATTCATACCGGGGTGAAACCTTATGGGTATCAGAAATGTGGAG 1440
QY 1441 AGCTTTAGTCAAGGCGATCAGCTTACACAAATCAGAAAAAGCAGAGTGGGCGAAATCC 1500
DB 1441 AGCTTTAGTCAAGGCGATCAGCTTACACAAATCAGAAAAAGCAGAGTGGGCGAAATCC 1500
QY 1501 TACGAATGTAAAGATGTGCGGAGAGCATGTAAACCATTAACATCTCCGAGAACATCAG 1560
DB 1501 TACGAATGTAAAGATGTGCGGAGAGCATGTAAACCATTAACATCTCCGAGAACATCAG 1560
QY 1561 AGGATCCACAAAGTGTGAAGAGCCTTTGTAAGAGAGTACCCGCTGTATCTATGTGTTTC 1620
DB 1561 AGGATCCACAAAGTGTGAAGAGCCTTTGTAAGAGAGTACCCGCTGTATCTATGTGTTTC 1620
QY 1621 GCTTTTCACAGTTTGTACCTGCAAGTCACTCAGTCAAAATATTAATGAAAAATTC 1680
DB 1621 GCTTTTCACAGTTTGTACCTGCAAGTCACTCAGTCAAAATATTAATGAAAAATTC 1680
QY 1681 CAGAAATTAAGAAATTTTAAGTCTCAATATGTGTGCCCTTCTAGATGAGCTGTGAATCT 1740
DB 1681 CAGAAATTAAGAAATTTTAAGTCTCAATATGTGTGCCCTTCTAGATGAGCTGTGAATCT 1740
QY 1741 CTGCGTGTCCGGCTCAGGCGGCGCGGAGATGTGATCATCCTTGTGTCCAGACATCCAC 1800
DB 1741 CTGCGTGTCCGGCTCAGGCGGCGCGGAGATGTGATCATCCTTGTGTCCAGACATCCAC 1800
QY 1801 GGTGTATAGCCACCGCTGCTATGTGATAGTACGCGTGTGTGTGTATCAGATCACT 1860
DB 1801 GGTGTATAGCCACCGCTGCTATGTGATAGTACGCGTGTGTGTGTATCAGATCACT 1860
QY 1861 ATCCAGACATCAAGTGTCTGTGCCAAGTATCTCACTTTGTCTTAAAGTGTGCCAG 1920
DB 1861 ATCCAGACATCAAGTGTCTGTGCCAAGTATCTCACTTTGTCTTAAAGTGTGCCAG 1920
QY 1921 AAGCAGAGATGATGTGCTGTGATTCGATATGCTCAAGAGAGAACCAAAAGTCTTC 1980
DB 1921 AAGCAGAGATGATGTGCTGTGATTCGATATGCTCAAGAGAGAACCAAAAGTCTTC 1980
```

Db 1921 AGAGCAGGAGTAGTGTCTGTGTATTCGGATATGCCAAGAGAGCCACCAAGTCTTC 1980
Qy 1981 CTTTAAATGAAAAGCTGAAAGTTCTCACTTAA 2015
Db 1981 CTTTAAATGAAAAGTGAAGTCTCAACTTAA 2015
RESULT 2
AAS03044
ID AAS03044 standard; cDNA, 2098 BP.
XX AAS03044;
AC AAS03044;
XX 29-AUG-2001 (first entry)
DE Human diagnostic and therapeutic (dithp) cDNA sequence #33.
XX
XX Human diagnostic and therapeutic molecule; dithp; gene therapy;
KW thalassemia; cardiovascular disorder; cell proliferative disorder;
KW cancer; neurodegenerative disorder; autoimmune disorder;
KW infectious disorder; inflammatory disorder; developmental disorder;
XX Incyte ID number 9027913dec; transcription factor molecule; ss.
XX Homo sapiens.
OS
XX WO200121836-A2.
PN
XX 29-MAR-2001.
PD
XX 19-SEP-2000; 2000MO-US025643.
PF
XX 23-SEP-1999; 99US-0155760P.
PR 24-SEP-1999; 99US-0155939P.
PR 24-SEP-1999; 99US-0156294P.
PR 28-SEP-1999; 99US-0156565P.
PR 28-SEP-1999; 99US-0156624P.
PR 28-SEP-1999; 99US-0156625P.
PR 24-NOV-1999; 99US-0167410P.
PR 24-NOV-1999; 99US-0167517P.
PR 24-NOV-1999; 99US-0167520P.
PR 24-NOV-1999; 99US-0167521P.
PR 24-NOV-1999; 99US-0167522P.
PR 24-NOV-1999; 99US-0167543P.
PR 24-NOV-1999; 99US-0167543P.
PR 29-NOV-1999; 99US-0167943P.
PR 29-NOV-1999; 99US-0167945P.
PR 30-NOV-1999; 99US-0168197P.
PR 30-NOV-1999; 99US-0168265P.
PR 30-NOV-1999; 99US-0168429P.
PR 30-NOV-1999; 99US-0168432P.
PR 01-DEC-1999; 99US-0168468P.
PR 01-DEC-1999; 99US-0168599P.
PR 02-DEC-1999; 99US-0168611P.
PR 02-DEC-1999; 99US-0168613P.
PR 02-DEC-1999; 99US-0168857P.
XX (INCY-) INCYTE GENOMICS INC.
XX
XX Hodgson DM, Lincoln SE, Russe PD, Spiro PA, Banville SC;
PI Bratcher SR, Duteour GE, Cohen HJ, Rosen BH, Shah P, Chalup MS;
PI Hillman JL, Jones AL, Yu JY, Greenwalt LB, Panzer SR, Rosberry AM;
PI Wright RJ, Chen W, Liu TF, Yap PE, Stockdreher TK, Amshay S,
PI Fong W;
XX WPI, 2001-281607/29.
XX
XX Novel diagnostic and therapeutic polynucleotides, used in disease
PT diagnosis and for gene therapy of conditions such as cancer and
PI thalassemia.
XX
XX Claim 1; Page 271; 299pp; English.
XX
XX The present sequence for human diagnostic and therapeutic (dithp) cDNA

CC sequence #33 is 1 of 71 (AAS03012-AAS03082) novel sequences described in
CC the invention. The present sequence (Incyte ID No: 9027913dec) encodes a
CC transcription factor molecule. The dithp polynucleotides may be used to
CC diagnose a condition disease or disorder associated with human molecules.
CC They can be used to identify the presence of similar nucleic acids. Dithp
CC polynucleotides may be used to generate hybridisation probes for use in
CC chromosomal mapping. Polypeptides (dithp) encoded by dithp are used to
CC screen for molecules which bind to them and modulate their activity.
CC Dithp polynucleotides can be used for gene therapy of disorders such as
CC severe combined immunodeficiency syndrome (SCID), cystic fibrosis,
CC thalassemia, haemophilia resulting from Factor VIII or IX deficiencies,
CC cardiovascular disorders e.g familial hypercholesterolaemia (FH), cell
CC proliferative disorders e.g. cancers, neurodegenerative disorders,
CC autoimmune/inflammatory disorders, infectious disorders and developmental
CC disorders. The antibodies can be used to analyse protein expression
CC levels
XX
SQ Sequence 2098 BP; 610 A; 467 C; 537 G; 484 T; 0 U; 0 Other;
Query Match 98.6%; Score 1986.8; DB 4; Length 2098;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 2010; Conservative 0; Mismatches 2; Indels 2; Gaps 2;
Qy 1 CGCTTCTGACCGGTGAGCAACCGTGTCTCGCCAGTCCGCGACGAGCATCCT 60
Db 85 CGCTTCTGACCGGTGAGCAACCGTGTCTCGCCAGTCCGCGACGAGCATCCT 144
Qy 61 TCAGAAAAGCATCCCGAGGAGGAGAGCAATGTTAAACATCTTAGTCACTTACG 120
Db 145 TCAGAAAAGCATCCCGAGGAGGAGAGCAATGTTAAACATCTTAGTCACTTACG 204
Qy 121 CTTTCGGAATTTGCTTTTTCAGTGGAAACCCCGAGAACTGATCTTTCAGTTCT 180
Db 205 CTTTCGGAATTTGCTTTTTCAGTGGAAACCCCGAGAACTGATCTTTCAGTTCT 264
Qy 181 AAAACAATGGCCAGGGGTTTGTGACGTTCCGCGACGATGACATGATCTTTCAGAG 240
Db 265 AAAACAATGGCCAGGGGTTTGTGACGTTCCGCGACGATGACATGATCTTTCAGAG 324
Qy 241 GAGTGGCGTGTCTGAACTCTCTGAGAGGACCTGTACTGGAGCTGATGCTGGAGAC 300
Db 325 GAGTGGCGTGTCTGAACTCTCTGAGAGGACCTGTACTGGAGCTGATGCTGGAGAC 384
Qy 301 TACAGTAATCTGTGCTCACTGATTTGGAGTCCAGATATGAAATTAAGTTTACTACA 360
Db 385 TACAGTAATCTGTGCTCACTGATTTGGAGTCCAGATATGAAATTAAGTTTACTACA 444
Qy 361 GAAAAAATCATTCATGAAATAGAGGCTCCAAAAGAAATTCAGATAGAAAGTAAATCC 420
Db 445 GAAAAAATCATTCATGAAATAGAGGCTCCAAAAGAAATTCAGATAGAAAGTAAATCC 504
Qy 421 CTGGCCGTAACTGATATGTGAAGAGTACGTTGAAGACCAACAGCGCTCAAGAGGAGG 480
Db 505 CTGGCCGTAACTGATATGTGAAGAGTACGTTGAAGACCAACAGCGCTCAAGAGGAGG 564
Qy 481 TATGTCATTCAGATGATATCAATTTATGTCAAAAGGCTGTCTACTAGGAAGGACCCCC 539
Db 565 TATGTCATTCAGATGATATCAATTTATGTCAAAAGGCTGTCTACTAGGAAGGACCCCC 624
Qy 540 TCCTAGAACACATCAGAGACATCATTAAGAGAAATTCCTTGAATGTAGAGCTGTGGAG 599
Db 625 TCCTAGAACACATCAGAGACATCATTAAGAGAAATTCCTTGAATGTAGAGCTGTGGAG 684
Qy 600 GGCCTTTAGTGTGTGATCAACTTAGTCAACATCAGAAAATTCATCTGTGAGAAAC 659
Db 685 GGCCTTTAGTGTGTGATCAACTTAGTCAACATCAGAAAATTCATCTGTGAGAAAC 744
Qy 660 TTATGAATGTAAAGAAATTAAGAGGCTTCGTTGGGCGAATAGCTTACTCAACATCA 719
Db 745 TTATGAATGTAAAGAAATTAAGAGGCTTCGTTGGGCGAATAGCTTACTCAACATCA 804
Qy 720 AAAAATTCATCTGAGGAGAGCCCTTACGAATGTAAAGATCTGTGGAGGCTTTTCGATG 779

```

Db      805 AAAAAATCACTACTGCGGAGAGAGCCCTACGAAATGTAAAGA CTGTGGAGAGGCTTTTCGATG 864
Qy      760 GGGCTCAAGCCTCGTTATTCATTAAGAGATTCATCT -GATGAAAAAACCTTAATGATGA 838
Db      865 GGGCTCAAGCCTCGTTATTCATTAAGAGATTCATCT GGGGTGAAAAAACCTTAATGATGA 924
Qy      839 AAGACTGTGGAAAGGCTTTGCGCGGTGTATGAGCTCA CTGAGCA CCGAGATTCACCA 898
Db      925 AAGACTGTGGAAAGGCTTTGCGCGGTGTATGAGCTCA CTGAGCA CCGAGATTCACCA 984
Qy      899 CTGGGGAGAAACATACGAAATGTGGGAAGACCTTTAGCCCTGTGTATTAAC 958
Db      985 CTGGGGAGAAACATACGAAATGTGGGAAGACCTTTAGCCCTGTGTATTAAC 1044
Qy      959 TTATTCAGACACAAGAAATTCATATGTGGGAGAGACCTTACGAGTGTAAAGA CTGTGGGA 1018
Db      1045 TTATTCAGACACAAGAAATTCATATGTGGGAGAGACCTTACGAGTGTAAAGA CTGTGGGA 1104
Qy      1019 AAGCTTTTATTTGTGGTTCAGGCTCATTCAGCATTAAGAAATTCACACAGGTGAGAAAC 1078
Db      1105 AAGCTTTTATTTGTGGTTCAGGCTCATTCAGCATTAAGAAATTCACACAGGTGAGAAAC 1164
Qy      1079 CCTATGAAGTGAAGAAATGTGGGAAGGCTTTACTGAGTCAATTCCTTTCTGACGATC 1138
Db      1165 CCTATGAAGTGAAGAAATGTGGGAAGGCTTTACTGAGTCAATTCCTTTCTGACGATC 1224
Qy      1139 AGAAGATCCACACCGGTGAGAAAGCCTCAGAAATGTAAAGATGTGGGAAGGCTTTGCGCT 1198
Db      1225 AGAAGATCCACACCGGTGAGAAAGCCTCAGAAATGTAAAGATGTGGGAAGGCTTTGCGCT 1284
Qy      1199 GGGGTTTCGAGCCTCGTTAAGCA CGAGAGATACATA CGGGCGAAGACCGGTACAGTGA 1258
Db      1285 GGGGTTTCGAGCCTCGTTAAGCA CGAGAGATACATA CGGGCGAAGACCGGTACAGTGA 1344
Qy      1259 CAGAAATGTGGGAAGGCTTCAATGTGGCTATTCACCTCACTCAGCA CCGAGAAATCCACA 1318
Db      1345 CAGAAATGTGGGAAGGCTTCAATGTGGCTATTCACCTCACTCAGCA CCGAGAAATCCACA 1404
Qy      1319 CAGGCGAAACCCCGTAAATGTAAAGATGTGGGAAGGCTTTCAATTTAGATCGAGCC 1378
Db      1405 CAGGCGAAACCCCGTAAATGTAAAGATGTGGGAAGGCTTTCAATTTAGATCGAGCC 1464
Qy      1379 TCGTGAACATGAGAGAAATTCATCCGGGGTGAACCTTAATGGGTGACAGAAATGTGGGA 1438
Db      1465 TCGTGAACATGAGAGAAATTCATCCGGGGTGAACCTTAATGGGTGACAGAAATGTGGGA 1524
Qy      1439 AAGACTTTAGTCAAGGCGCATCAGCTTACACACATACAGAAAAAGCACAGTGGGGCGAAAT 1498
Db      1525 AAGACTTTAGTCAAGGCGCATCAGCTTACACACATACAGAAAAAGCACAGTGGGGCGAAAT 1584
Qy      1499 CCTACGAATGTAAAGAGTCCGGGAAGGCAATGTAAACCACTTAACCACTTCCGAGAACATC 1558
Db      1585 CCTACGAATGTAAAGAGTCCGGGAAGGCAATGTAAACCACTTCCGAGAACATC 1644
Qy      1559 AAGAGATCCACACAGTGTGAAGGCTTTTGAACGAGTAGCCGCTCGTATCTATGTT 1618
Db      1645 AAGAGATCCACACAGTGTGAAGGCTTTTGAACGAGTAGCCGCTCGTATCTATGTT 1704
Qy      1619 TCGCTTTCCACAGTTTGTATCCCTGAGTCAACCTGCAAGTTCAAAAAATTTAAATGAAAAAT 1678
Db      1705 TCGCTTTCCACAGTTTGTATCCCTGAGTCAACCTGCAAGTTCAAAAAATTTAAATGAAAAAT 1764
Qy      1679 TCCAGAAATTAAGAAATTTTAAGTCTGAAGATGTGTGCGCTTCTGAGTAGCGGTATGAAT 1738
Db      1765 TCCAGAAATTAAGAAATTTTAAGTCTGAAGATGTGTGCGCTTCTGAGTAGCGGTATGAAT 1824
Qy      1739 CTCTCGCTGTCCGGCTCCAGCCGCGCGGGATGTAGTCAATCCCTTGTGTCACACATCC 1798
Db      1825 CTCTCGCTGTCCGGCTCCAGCCGCGCGGGATGTAGTCAATCCCTTGTGTCACACATCC 1884
Qy      1799 AGGCTGTATACGCCACCCACCTGCTAGTACTAGTACCGCGCTTGGTGTATCAATCAA 1858
Db      1885 AGGCTGTATACGCCACCCACCTGCTAGTACTAGTACCGCGCTTGGTGTATCAATCAA 1944

```

```

Qy      1859 CTATCCAGCATCAGTGCCTGTGCCCCCAAGTAGTCTCACTTTGCTTAAACAGTGGCCCC 1918
Db      1945 CTATCCAGCATCAGTGCCTGTGCCCCCAAGTAGTCTCACTTTGCTTAAACAGTGGCCCC 2004
Qy      1919 AGAGAGCAGAGAGTGAATGCTGTGTGATTCGGATATGCCAAAGAGAGAGCCACAAAGTGCT 1978
Db      2005 AGAGAGCAGAGAGTGAATGCTGTGTGATTCGGATATGCCAAAGAGAGAGCCACAAAGTGCT 2064
Qy      1979 TCCTTTAAATGAAAGGAGGAAGTTCCAACTT 2012
Db      2065 TCCTTTAAATGAAAGGAGGAAGTTCCAACTT 2098

RESULT 3
ABA82974
ID   ABA82974 standard; DNA; 2173 BP.
XX
XX   ABA82974;
AC
XX
XX   05-FEB-2002 (first entry)
DT
XX
DE   Human transcription factor TRFX-1 coding sequence.
XX
XX   Human; transcription factor; TRFX; cell proliferative disease;
KW   autoimmune disease; inflammation; neurological disease;
KW   developmental disorder; cancer; AIDS; infection; cytostatic; anti-HIV;
KW   neuroprotective; antiinflammatory; gene therapy; db.
XX
XX   Homo sapiens.
OS
XX   WO200172777-A2.
PN
XX
XX   04-OCT-2001.
PD
XX
XX   13-MAR-2001; 2001WO-US008117.
PF
XX
XX   13-MAR-2000; 2000US-0186986P.
PR
XX
XX   (INCYTE) INCYTE GENOMICS INC.
PA
XX
XX   Hillman JL, Baughn WR, Yue H, Lal P, Lu DM, Patterson C,
PI   Azimzal Y, Bandman O, Tang YT, Matchur P, Shah P, Au-Young J,
PI   Reddy R;
XX
XX   MPI: 2001-570896/64.
DR   P-PSDB; ABB50150.
DR
XX
XX   Novel transcription factor polypeptides, used to treat diseases
PT   associated with altered activity and expression of TRFX, and to screen
PT   for agents capable of modulating its activity.
XX
XX   Claim 11; Page 252; 327pp; English.
PS
XX
XX   The present sequence is the coding sequence for a human transcription
CC   factor. The transcription factor and its coding sequence are useful in
CC   the diagnosis, treatment and prevention of diseases associated with
CC   altered expression of the transcription factor e.g. cell proliferative,
CC   autoimmune/inflammatory, neurological and developmental disorders. A
CC   number of specific disorders/diseases are given in the specification,
CC   including: arteriosclerosis, cirrhosis, hepatitis, cancers, AIDS,
CC   allergies, anaemia, asthma, autoimmune thyroiditis, bronchitis, atopic
CC   dermatitis, diabetes mellitus, emphysema, Goodpasture's syndrome, gout,
CC   Grave's disease, multiple sclerosis, osteoarthritis, pancreatitis,
CC   psoriasis, rheumatoid arthritis, systemic lupus erythematosus, ulcerative
CC   colitis, uveitis, Alzheimer's disease, Huntington's disease, Parkinson's
CC   disease, stroke, and viral, bacterial, fungal and protozoal infections
CC
XX
XX   Sequence 2173 BP; 648 A; 482 C; 555 G; 488 T; 0 U; 0 Other;

Query Match      97.6%; Score 1967.4; DB 5; Length 2173;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1968; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

47 CAGGCCAGATCCTTCAGAAAAAGCATCCCGAGGAGGAAGACGATGTTAAACATCTT 106
Db CAGGCCAGATCCTTCAGAAAAAGCATCCCGAGGAGGAAGACGATGTTAAACATCTT 248
Qy AGGTGAGCTTAGCCTCTCGAATTTGTTCTTCACTGGAACCCCGAAGACTGATC 166
Db AGGTGAGCTTAGCCTCTCGAATTTGTTCTTCACTGGAACCCCGAAGACTGATC 308
Qy AGTTCTTCACTTCAAAAACATGCGCCAGGCTTGTGACGTTCCCGCAGCTAGCCATAG 226
Db AGTTCTTCACTTCAAAAACATGCGCCAGGCTTGTGACGTTCCCGCAGCTAGCCATAG 368
Qy ACTTTCTCAGGAGAGTGGGCTGTGAACTCTGACCTCAGAGGACCTGTGCTGGAGC 286
Db ACTTTCTCAGGAGAGTGGGCTGTGAACTCTGACCTCAGAGGACCTGTGCTGGAGC 428
Qy TGATGCTGAGAACTACAGTAACCTTGCTCACTGGAATTTGGAAGTCAAGATAGAAATA 346
Db TGATGCTGAGAACTACAGTAACCTTGCTCACTGGAATTTGGAAGTCAAGATAGAAATA 488
Qy AGAGTTTACCTACAGAAAAAAACATTCATGAATAAGGCTTCCAAAAGAAATTCAGATA 406
Db AGAGTTTACCTACAGAAAAAAACATTCATGAATAAGGCTTCCAAAAGAAATTCAGATA 548
Qy GAAGAAGTAAATCCCTTGGCCGTAATGGAATGGAAGTACGCTTGAAGACCAACAGC 466
Db GAAGAAGTAAATCCCTTGGCCGTAATGGAATGGAAGTACGCTTGAAGACCAACAGC 608
Qy GCTCAGAGGAGGATGATCATCATGATGATCATCAATTAATGTCAAAAGGCTGCTACTA 526
Db GCTCAGAGGAGGATGATGATCATCATGATGATCATCAATTAATGTCAAAAGGCTGCTACTA 668
Qy GAGAAAGCACCCTCTCTAGAACATCATCAGAGACATCATPAAGGAAATTCCTTGAATGTA 586
Db GAGAAAGCACCCTCTCTAGAACATCATCAGAGACATCATPAAGGAAATTCCTTGAATGTA 728
Qy AGGATCTGGGAGAGGCTTTAGTGTGGCTATCACTTAATGTAACATCAAGAAATTCCTA 646
Db AGGATCTGGGAGAGGCTTTAGTGTGGCTATCACTTAATGTAACATCAAGAAATTCCTA 788
Qy CTGTGAGAAACCTTATGATGTAAGATGTAAGAGGCTTCCGTTGGGCAATCAAGC 706
Db CTGTGAGAAACCTTATGATGTAAGATGTAAGAGGCTTCCGTTGGGCAATCAAGC 848
Qy TTACTCAACATCAAAAAATTCATCTGAGGAGAGCCCTACGAATGTAAAGACTGTGGA 766
Db TTACTCAACATCAAAAAATTCATCTGAGGAGAGCCCTACGAATGTAAAGACTGTGGA 908
Qy AGGCTTTTCGATGGGCTCAAGCCTCGTTATTCATPAAGAGATTCATCTGTGTAAGAAC 826
Db AGGCTTTTCGATGGGCTCAAGCCTCGTTATTCATPAAGAGATTCATCTGTGTAAGAAC 968
Qy CCTATGATGTAAAGACTGTGGAAGGCTTTCGCGTGTGTGATGAGCTCACTCAGACCC 886
Db CCTATGATGTAAAGACTGTGGAAGGCTTTCGCGTGTGTGATGAGCTCACTCAGACCC 1028
Qy AGAGATTCACAACCTGGGAGAAAGACTACGAATGTGAAGACTGTGGAAAGACTTTAAGC 946
Db AGAGATTCACAACCTGGGAGAAAGACTACGAATGTGAAGACTGTGGAAAGACTTTAAGC 1088
Qy GTGTGTAATTAATCTTTCAGCAACAGAAATTCATGATGGGAGAAAGCTTACAGATGTA 1006
Db GTGTGTAATTAATCTTTCAGCAACAGAAATTCATGATGGGAGAAAGCTTACAGATGTA 1148
Qy AAGACTGTGGAAAGGCTTTTATTTGTTCAAGCTCATTCAGCATTAAGAAATTCACA 1066
Db AAGACTGTGGAAAGGCTTTTATTTGTTCAAGCTCATTCAGCATTAAGAAATTCACA 1208
Qy CAGGTGAAGAAACCTTATGATGTCAAGATGTGGAGAGGCTTTTCTGAGTCAATTAAC 1126
Db CAGGTGAAGAAACCTTATGATGTCAAGATGTGGAGAGGCTTTTCTGAGTCAATTAAC 1268

Qy TTACTCAGATCAGAAAGATCCACACCGGTGAGAAAGCTCACAAGATGTAAAGAGTGGGA 1186
Db TTACTCAGATCAGAAAGATCCACACCGGTGAGAAAGCTCACAAGATGTAAAGAGTGGGA 1328
Qy AGGCTTTTCGCTGGGCTTCAGGCTCGTTAAGCAGAGAGATACATACGGGCGAAGAC 1246
Db AGGCTTTTCGCTGGGCTTCAGGCTCGTTAAGCAGAGAGATACATACGGGCGAAGAC 1388
Qy CGTACAGTGCACAGAAATGTGGAAAGGCTTCAATTTGTGCTATTAATCTCACTCAGACG 1306
Db CGTACAGTGCACAGAAATGTGGAAAGGCTTCAATTTGTGCTATTAATCTCACTCAGACG 1448
Qy AGAGATTCACAACAGGCGAAACCCCGTATTAATGTAAAGAGTGGGAAAGCTTCAATT 1366
Db AGAGATTCACAACAGGCGAAACCCCGTATTAATGTAAAGAGTGGGAAAGCTTCAATT 1508
Qy ATGATCAGGCTCGTGAACATGAGAAATTCATACCGGGTGAACCTTATGGGTGTA 1426
Db ATGATCAGGCTCGTGAACATGAGAAATTCATACCGGGTGAACCTTATGGGTGTA 1568
Qy CAGAAATGTGGAAAGCTTTAGTCAAGGCTCATGAGCTTACATCAATCAGAAAAAGCACA 1486
Db CAGAAATGTGGAAAGCTTTAGTCAAGGCTCATGAGCTTACATCAATCAGAAAAAGCACA 1628
Qy GTGGGGGGAATCCTACGAATGTAAAGAGTGGGAGGAGCANTGTAAACACCTTAACATC 1546
Db GTGGGGGGAATCCTACGAATGTAAAGAGTGGGAGGAGCANTGTAAACACCTTAACATC 1688
Qy TCCGAGAAACATCAGAGATCCACAACAGTTGAAGAGCTTTTGAACGAGTAGCCGCTC 1606
Db TCCGAGAAACATCAGAGATCCACAACAGTTGAAGAGCTTTTGAACGAGTAGCCGCTC 1748
Qy GTATCTATGTTTGGCTTCCACAGTTTGTACCTGCACTGCACTGCAATTAAGAAATAT 1666
Db GTATCTATGTTTGGCTTCCACAGTTTGTACCTGCACTGCACTGCAATTAAGAAATAT 1808
Qy TAAATGAAAAATTCAGAAATTAAGAAATTTAAGTCTAATAATGTGTGCCCTTTCAGATA 1726
Db TAAATGAAAAATTCAGAAATTAAGAAATTTAAGTCTAATAATGTGTGCCCTTTCAGATA 1868
Qy GCGTATGAAATCTCTCGCTGTCCGCTCCAGCCGCGCGGGAGTGTAGTCAATCCCTTGG 1786
Db GCGTATGAAATCTCTCGCTGTCCGCTCCAGCCGCGCGGGAGTGTAGTCAATCCCTTGG 1928
Qy TCCAGCACATCCAGCTGTATACGCCACCCAGCTGTAGTGAAGTGTAGGCGCTTGG 1846
Db TCCAGCACATCCAGCTGTATACGCCACCCAGCTGTAGTGAAGTGTAGGCGCTTGG 1988
Qy TGATCAGATCAACTATCCAGCATCAGAGTGTGCTGCAAGTGTGTCTCACTTGGCTT 1906
Db TGATCAGATCAACTATCCAGCATCAGAGTGTGCTGCAAGTGTGTGTCTCACTTGGCTT 2048
Qy AACAGTGGCCCAAGAGACAGAGTGTGATGCTGTGATTCGATATGCAAGAGAG 1966
Db AACAGTGGCCCAAGAGACAGAGTGTGATGCTGTGATTCGATATGCAAGAGAG 2108
Qy CCACAAAGTGTCTCTTTTAAATGAAAAAGGTGAAGTTCTCACTTAA 2015
Db CCACAAAGTGTCTCTTTTAAATGAAAAAGGTGAAGTTCTCACTTAA 2157

RESULT 4
AAK94562 standard; cDNA; 2169 BP.
AAK94562;
AC AAK94562;
DT 06-NOV-2001 (first entry)
XX Human full-length cDNA, SEQ ID NO: 3469.
DE Human; full length cDNA; cDNA synthesis; oligo-capping; ss.
XX
XX

OS Homo sapiens.
 XX EPI130094-A2.
 XX
 XX 05-SEP-2001.
 XX
 PF 07-JUL-2000; 2000EP-00114089.
 XX
 XX 08-JUL-1999; 99JP-00194486.
 PR 11-JAN-2000; 2000JP-00118774.
 PR 02-MAY-2000; 2000JP-00183765.
 XX
 PA (HELI-) HELIX RES INST.
 XX
 PI Ota T, Nishikawa T, Isegai T, Hayaishi K, Ishii S, Kawai Y,
 PI Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;
 XX
 XX WPI; 2001-524255/58.
 DR P-PSDB; AAM93629.
 XX
 XX 830 Primers useful for synthesizing full length cDNA clones and their use
 PT in genetic manipulation.
 XX
 PS Claim 8; SEQ ID NO 3469; 1380bp + Sequence Listing; English.
 XX
 CC The invention relates to primers for synthesizing full length cDNA
 CC clones. 830 cDNA molecules encoding a human protein have been isolated
 CC and nucleotide sequences of 5'- and 3'-ends of the cDNA molecules have
 CC been determined. Primers for synthesizing the full length cDNA are useful
 CC for clarifying the function of the protein encoded by the cDNA. The full
 CC length clones were obtained by construction of full length enriched cDNA
 CC libraries that were synthesised by the oligo-capping method. The primers
 CC enable the production of the full length cDNA easily without any special
 CC methods. The present sequence is a full length human cDNA of the
 CC invention. Note: The sequence data for this patent did not form part of
 CC the printed specification, but was obtained in CD-ROM format directly
 CC from EPO
 XX
 SQ Sequence 2169 BP; 633 A; 485 C; 560 G; 491 T; 0 U; 0 Other;

Query Match 97.2%; Score 1958; DB 4; Length 2169;
 Best Local Similarity 99.7%; Pred. No. 0;
 Matches 1961; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 47 CAGGCCAGCATCTCTGAGAAAAAGCATCCCGAGAGAGAAAGCAATCGTTAAACATCTT 106
 DB 204 CAGGCCAGCATCTCTGAGAAAAAGCATCCCGAGAGAGAAAGCAATCGTTAAACATCTT 263
 QY 107 AGGTGAGCTCTAGCCCTCTCGGAAATTTGTCTTTCAGTGAGAAACCCCGAAGACATGATC 166
 DB 264 AGGTGAGCTCTAGCCCTCTCGGAAATTTGTCTTTCAGTGAGAAACCCCGAAGACATGATC 323
 QY 167 AGTCTTCAGTTTAAACAATGAGGCCAGGGTTTGTGACGTTGCCGAGAGTACGCAATG 226
 DB 324 AGTCTTCAGTTTAAACAATGAGGCCAGGGTTTGTGACGTTGCCGAGAGTACGCAATG 383
 QY 227 ACTTTCTGAGAGAGTGGGCTGTCTGAACTCTGCTGAGAGGACCTGTACTGGGAGG 286
 DB 384 ACTTTCTGAGAGAGTGGGCTGTCTGAACTCTGCTGAGAGGACCTGTACTGGGAGG 443
 QY 287 TGAATGAGAACTACAGTAATCTGATCTCACTGAGATTTGAGATGACATATGAAATA 346
 DB 444 TGAATGAGAACTACAGTAATCTGATCTCACTGAGATTTGAGATGACATATGAAATA 503
 QY 347 AGAGTTTACCTTAAGAAAAAAACATTCATGAATTAAGGGCTTCCAAAGAGAAATTCAGAT 406
 DB 504 AGAGTTTACCTTAAGAAAAAAACATTCATGAATTAAGGGCTTCCAAAGAGAAATTCAGAT 563
 QY 407 GAAGAAGTAATCTCTGGCCGTAATGATGATGAGAGTACGTTGAAAGCCACAGC 466
 DB 564 GAAGAAGTAATCTCTGGCCGTAATGATGATGAGAGTACGTTGAAAGCCACAGC 623
 QY 467 GCTCCAGAGGAGTATGTCATCAGATGATCATCAATTATGTCAAAAGGCTGTACTA 526

DB 624 GCTCCAGAGGAGGAGTATGTCATCAGATGATCATCAATTATGTCAAAAGGCTGTACTA 683
 QY 527 GAGAAGGACCCCTCTGAAACACATCAGAGCATCATAGGAAATTCCTTGAATGTA 586
 DB 684 GAGAAGGACCCCTCTGAAACACATCAGAGCATCATAGGAAATTCCTTGAATGTA 743
 QY 587 AGGACTGGGAAAGGCTTTAGTCGTGGCTATCACTTAGTCAACATCAGAAAAATCCATA 646
 DB 744 AGGACTGGGAAAGGCTTTAGTCGTGGCTATCACTTAGTCAACATCAGAAAAATCCATA 803
 QY 647 CTGTGAGAAACCTTATGATATGTAAGATGTAAGAGGCTTCGCTGGGGCAATCAGC 706
 DB 804 CTGTGAGAAACCTTATGATATGTAAGATGTAAGAGGCTTCGCTGGGGCAATCAGC 863
 QY 707 TTACTCAACATCAAAAAAATTCATCTGGGGAGAACCCCTAGAGATGTAAGACTGAGGA 766
 DB 864 TTACTCAACATCAAAAAAATTCATCTGGGGAGAACCCCTAGAGATGTAAGACTGAGGA 923
 QY 767 AGGCTTTTCGATGGGGCTCAAGCCTCGTTATTCATAGAGATTCATCTGGTGAANAAC 826
 DB 924 AGGCTTTTCGATGGGGCTCAAGCCTCGTTATTCATAGAGATTCATCTGGTGAANAAC 983
 QY 827 CCTATGATGTAAAGACTGTGGAAGGCTTTTCGGCGTGTGATGAGCTCACTCAGCACC 886
 DB 984 CCTATGATGTAAAGACTGTGGAAGGCTTTTCGGCGTGTGATGAGCTCACTCAGCACC 1043
 QY 887 AGAGATTCACACTGGGGGAGAAAGCTAGCAATGCAAGACCTGTGGAGAACCTTTAGCC 946
 DB 1044 AGAGATTCACACTGGGGGAGAAAGCTAGCAATGCAAGACCTGTGGAGAACCTTTAGCC 1103
 QY 947 GTGTGTAAATCTTATTTAGCAACAAGAAATTCATAGTGGGAGAGCCTTACGAGTGA 1006
 DB 1104 GTGTGTAAATCTTATTTAGCAACAAGAAATTCATAGTGGGAGAGCCTTACGAGTGA 1163
 QY 1007 AAGACTGTGGAAAGCCTTTATTTGTGTTCAAGCTCTATTAGCATTAAGAAATTCACA 1066
 DB 1164 AAGACTGTGGAAAGCCTTTATTTGTGTTCAAGCTCTATTAGCATTAAGAAATTCACA 1223
 QY 1067 CAGGTGAGAAACCCATATGATGTCAGAAATGAGGAGGCTTTACTGAGATCAATTACC 1126
 DB 1224 CAGGTGAGAAACCCATATGATGTCAGAAATGAGGAGGCTTTACTGAGATCAATTACC 1283
 QY 1127 TTACTCAGCATCAGAAAGATCCACACCGGTGAGAAAGCTTACGAAATGTAAGAGTGGGA 1186
 DB 1284 TTACTCAGCATCAGAAAGATCCACACCGGTGAGAAAGCTTACGAAATGTAAGAGTGGGA 1343
 QY 1187 AGGCTTTTCGTGGGGTTGAGACCTTCGTTAAGCAGAGAGATCACTAGGGCGAGAAC 1246
 DB 1344 AGGCTTTTCGTGGGGTTGAGACCTTCGTTAAGCAGAGAGATCACTAGGGCGAGAAC 1403
 QY 1247 CGTACAGTGCACAGAAATGTGGGAAAGGCTTCAATTTGTGGCTATCACTCAGACAG 1306
 DB 1404 CGTACAGTGCACAGAAATGTGGGAAAGGCTTCAATTTGTGGCTATCACTCAGACAG 1463
 QY 1307 AGAGAAATCCACACAGGCGAAACCCCGTATTAATGTAAGAGTGTGGGAAAGGCTTCAATT 1366
 DB 1464 AGAGAAATCCACACAGGCGAAACCCCGTATTAATGTAAGAGTGTGGGAAAGGCTTCAATT 1523
 QY 1427 CAGAAATGTGGGAAAGGCTTTAATCCCGGCATCAGCTTACACAAACATCAGAAAAAGCA 1486
 DB 1524 ATGGAATCGAGCCTCGTGTAACTGAGAGAAATTCATACCGGGGTGAAACCTTATGGGTGTA 1583
 QY 1427 CAGAAATGTGGGAAAGGCTTTAATCCCGGCATCAGCTTACACAAACATCAGAAAAAGCA 1486
 DB 1584 CAGAAATGTGGGAAAGGCTTTAATCCCGGCATCAGCTTACACAAACATCAGAAAAAGCA 1643
 QY 1487 GTGGGGCGAAATCTTAAGAAATGTAAGAGTGTGGGAAAGGCAATGTAACCACTTAACATC 1546
 DB 1644 GTGGGGCGAAATCTTAAGAAATGTAAGAGTGTGGGAAAGGCAATGTAACCACTTAACATC 1703
 QY 1547 TCCGAGAACATCAGAGATCCAAACAGTTGAAGAGCTTTTGAACGAGTAGCCGCTC 1606

Db 1704 TCCGAGACATGACGAGATCCACAAAGTGAAGAGCCTTTTGAAGCAGTAGACCCGCTC 1763
 QY 1607 GATATCATAGTGTTCGTTCCACAGTTTGTACTTCAGTCAATCCAGTTCAAAATAT 1666
 Db 1764 GATATCATAGTGTTCGTTCCACAGTTTGTACTTCAGTCAATCCAGTTCAAAATAT 1823
 QY 1667 TAAATGAAAAATTCAGAAATAAAGATTTTAAGTCTCAAAAGTGTGCTTCTGAGTA 1726
 Db 1824 TAAATGAAAAATTCAGAAATAAAGATTTTAAGTCTCAAAAGTGTGCTTCTGAGTA 1883
 QY 1727 GCGTGAATGAATCTCTCGCTGTCCGCTCCAGCCGCGGAGTGTAGTCAATCCCTTG 1786
 Db 1884 GCGTGAATGAATCTCTCGCTGTCCGCTCCAGCCGCGGAGTGTAGTCAATCCCTTG 1943
 QY 1787 TCCAGACATCCACGCTGTATACGCCACCACTTGTCTAGTGAATGAGCCGCTTG 1846
 Db 1944 TCCAGACATCCACGCTGTATACGCCACCACTTGTCTAGTGAATGAGCCGCTTG 2003
 QY 1847 TGAATGATCAATCAATCCAGATCCAGATCCAGTGCCTGTGCCAAGTACTCTCTGCTT 1906
 Db 2004 TGAATGATCAATCAATCCAGATCCAGATCCAGTGCCTGTGCCAAGTACTCTCTGCTT 2063
 QY 1907 AACAGTGGCCCAAGAGAGAGAGTAGTAGTGTGTGTGATTCGATATGCCAAGAAG 1966
 Db 2064 AACAGTGGCCCAAGAGAGAGAGTAGTAGTGTGTGTGATTCGATATGCCAAGAAG 2123
 QY 1967 CCACAAAGTCTTCTCTTTTAAATGAAAAGTGAAGTCTCACTT 2012
 Db 2124 CCACAAAGTCTTCTCTTTTAAATGAAAAGTGAAGTCTCACTT 2169

RESULT 5

ADL31436
ID ADL31436 standard; cDNA; 2169 BP.

ADL31436;

20-MAY-2004 (first entry)

Full length human cDNA clone Segid 3469.

human; medicine; signal transduction; glycoprotein; transcription;
oligo-capping method; seq. gene.

Homo sapiens.

EP1396543-A2.

10-MAR-2004.

07-JUL-2000; 2003EP-00025638.

08-JUL-1999; 99JP-00194486.

11-JAN-2000; 2000JP-00118774.

02-MAY-2000; 2000JP-00183865.

07-JUL-2000; 2000EP-00114089.

(REAS-) RES ASSOC BIOTECHNOLOGY.

Ota T, Nishikawa T, Isegai T, Hayashi K, Ishii S, Kawai Y;
Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;
F-PSDB; ADL31437.

New oligonucleotide primers (830 cDNAs) useful for synthesizing full
length human cDNAs.

Example 1; SEQ ID NO 3469; 1340bp, English.

This invention relates to a novel primers useful for synthesizing full
length cDNA molecules that encode human proteins. Specifically, it refers
to secretory or membrane proteins that are potential therapeutic agents/

CC target molecules in the field of medicine, and in particular genes
CC encoding proteins that are associated with signal transduction,
CC glycoproteins and transcription. The present invention describes a method
CC for efficiently cloning a full length human cDNA from both the 5' and 3'
CC ends using the oligo-capping method. This polynucleotide sequence is a
CC full length human cDNA clone of the invention.

XX Sequence 2169 BP; 633 A; 485 C; 560 G; 491 T; 0 U; 0 Other;

Query Match 97.2%; Score 1958; DB 12; Length 2169;

Best Local Similarity 99.7%; Pred. No. 0;

Matches 1961; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 47 CAGGCCAGATCTCTTCAAAAAAGCATCCCGAGAGAGAAAGCAATGCTTAAACTTT 106
 Db 204 CAGGCCAGATCTCTTCAAAAAAGCATCCCGAGAGAGAAAGCAATGCTTAAACTTT 263
 QY 107 AGGTGAGCTCTAGCTCTCGAATTTTGTCTTCTTCAAGTGAAGAACCCGAGAGACTGATC 166
 Db 264 AGGTGAGCTCTAGCTCTCGAATTTTGTCTTCTTCAAGTGAAGAACCCGAGAGACTGATC 323
 QY 167 AGTCTTCAGTTCTAAAACAATGAGCCAGAGTTTGTGTGACGTTGCGCAGAGTACCATAG 226
 Db 324 AGTCTTCAGTTCTAAAACAATGAGCCAGAGTTTGTGTGACGTTGCGCAGAGTACCATAG 383
 QY 227 ACTTTTCTCAGAGAGAGTGGGCTGTCTGAACTCTGCTCAGAGGAGCCTGTACTGGGAGC 286
 Db 384 ACTTTTCTCAGAGAGAGTGGGCTGTCTGAACTCTGCTCAGAGGAGCCTGTACTGGGAGC 443
 QY 287 TGATGCTGGAGAACTACAGTAATCTGTGCTCACTGGAATTTGGAGTCAGCATATGAAATA 346
 Db 444 TGATGCTGGAGAACTACAGTAATCTGTGCTCACTGGAATTTGGAGTCAGCATATGAAATA 503
 QY 347 AGAGTTTACCTACAGAAAAAAACATTCAATGAATAAAGGCTTCAAAAGAAATTCAGATA 406
 Db 504 AGAGTTTACCTACAGAAAAAAACATTCAATGAATAAAGGCTTCAAAAGAAATTCAGATA 563
 QY 407 GAAAGAAATTAATCCCTTGGCCGTAACTGATATGTGAAGATACGCTTGAAAGAACCAACAGC 466
 Db 564 GAAAGAAATTAATCCCTTGGCCGTAACTGATATGTGAAGATACGCTTGAAAGAACCAACAGC 623
 QY 467 GCTCCAGAGGAGAGTAGATGCAATCAGATGATCATCAATATATGCAAAAGACCTGTACTTA 526
 Db 624 GCTCCAGAGGAGAGTAGATGCAATCAGATGATCATCAATATATGCAAAAGACCTGTACTTA 683
 QY 527 GAGAAAGCAACCCCTCTAGAAACATCAAGACATCATTAAGAGAAATTCCTTTGAATGTA 586
 Db 684 GAGAAAGCAACCCCTCTAGAAACATCAAGACATCATTAAGAGAAATTCCTTTGAATGTA 743
 QY 587 AGGACTGTGGGAAAGGCTTTAGTGTGTGTATCACTTAATGTAAGTCAAGAAATTCATTA 646
 Db 744 AGGACTGTGGGAAAGGCTTTAGTGTGTGTATCACTTAATGTAAGTCAAGAAATTCATTA 803
 QY 647 CTGTGAGAGAACTTATGTAATGTAAGAATGTGAAGAGGCTTCCGTTGGGGCAATCAGC 706
 Db 804 CTGTGAGAGAACTTATGTAATGTAAGAATGTGAAGAGGCTTCCGTTGGGGCAATCAGC 863
 QY 707 TTAATCAACATCAAAAAATTTACTGTGGGAGAGACCTTACAGATGTAAAGACTGTGGGA 766
 Db 864 TTAATCAACATCAAAAAATTTACTGTGGGAGAGACCTTACAGATGTAAAGACTGTGGGA 923
 QY 767 AGGCTTTTCTGATGGGCTCAGACCTGTATTCATTAAGAGATTCATTAAGTGAAGAAAC 826
 Db 924 AGGCTTTTCTGATGGGCTCAGACCTGTATTCATTAAGAGATTCATTAAGTGAAGAAAC 983
 QY 827 CCTATGAATGTAAAGACTGTGAAGAGGCTTCCGCGTGTGATGAGCTCATCAGACACC 886
 Db 984 CCTATGAATGTAAAGACTGTGAAGAGGCTTCCGCGTGTGATGAGCTCATCAGACACC 1043
 QY 887 AGAGATTCACACTGGGAGAAAGACTTAAGATGCAAAAGACTGTGGGAAAGACTTTAGCC 946
 Db 1044 AGAGATTCACACTGGGAGAAAGACTTAAGATGCAAAAGACTGTGGGAAAGACTTTAGCC 1103

947 GTGTGATTAACCTATTATTCAGACACAGAAATTCATATGAGGAGAGACCTTACGAGTGA 1006
1104 GTGTGATTAACCTATTATTCAGACACAGAGATTCATATGAGGAGAGACCTTACGAGTGA 1163
1007 AAGACTGTGGGAAAGGCTTTTATTGTGTTCAAGCCTCATTCAGCATTAAGAAATTCACA 1066
1164 AAGACTGTGGGAAAGGCTTTTATTGTGTTCAAGCCTCATTCAGCATTAAGAAATTCACA 1223
1067 CAGGTGAGAAACCTTGAATGTCAGAAATGTGGGAAAGGCTTTTACCTGAGTGAATTAAC 1126
1224 CAGGTGAGAAACCTTGAATGTCAGAAATGTGGGAAAGGCTTTTACCTGAGTGAATTAAC 1283
1127 TTAATCAGCATCAGAAATTCACACCGGTGAGAGCCTTCAAGATGTAAGAGTGGGA 1186
1284 TTAATCAGCATCAGAAATTCACACCGGTGAGAGCCTTCAAGATGTAAGAGTGGGA 1343
1187 AGGCTTTTGGCTGGGTTGAGCTTGTTAAGCAGAGAGATACATACGGGGAGAGAGC 1246
1344 AGGCTTTTGGCTGGGTTGAGCTTGTTAAGCAGAGAGATACATACGGGGAGAGAGC 1403
1247 CGTACAGTGCACAGAAATGTGGGAAAGGCTTTCAATTGTGGCTATCACTTCACTCAGACG 1306
1404 CGTACAGTGCACAGAAATGTGGGAAAGGCTTTCAATTGTGGCTATCACTTCACTCAGACG 1463
1307 AAGAAATCCACACAGGCGAAACCCCGTAAATGTAAAGAGTGTGGGAAAGCTTTCAATT 1366
1464 AAGAAATCCACACAGGCGAAACCCCGTAAATGTAAAGAGTGTGGGAAAGCTTTCAATT 1523
1367 ATGATCGAGCCTCGTGAACATGAGAGAAATTCATACCGGGGGTGAACCTTATGGTGA 1426
1524 ATGATCGAGCCTCGTGAACATGAGAGAAATTCATACCGGGGGTGAACCTTATGGTGA 1583
1427 CAGAAATGTGGGAAAGGCTTTAGTCACGGCCATCAGCTTACACAAATCAGAAACGCA 1486
1584 CAGAAATGTGGGAAAGGCTTTAGTCACGGCCATCAGCTTACACAAATCAGAAACGCA 1643
1487 GTGGGGCGAAATCTTGAATGTAAAGAGTGTGGGAAAGGCTTAAACCTTAACCATC 1546
1644 GTGGGGCGAAATCTTGAATGTAAAGAGTGTGGGAAAGGCTTAAACCTTAACCATC 1703
1547 TCCGAAACATCAGAGATTCACAAAGTGAAGAGCCTTTTGAAGCAGTACCCGCTC 1606
1704 TCCGAAACATCAGAGATTCACAAAGTGAAGAGCCTTTTGAAGCAGTACCCGCTC 1763
1607 GATCTATGCTTCCGCTTCCACAGTTTGTACCTGAGTCAACTGAGTTCAAAAATAT 1666
1764 GATCTATGCTTCCGCTTCCACAGTTTGTACCTGAGTCAACTGAGTTCAAAAATAT 1823
1667 TAAATGGAATAATTCAGAAATAAGAAATTTAAAGTCTCAATGAGTGTGCCCTTCTGAGTA 1726
1824 TAAATGGAATAATTCAGAAATAAGAAATTTAAAGTCTCAATGAGTGTGCCCTTCTGAGTA 1883
1727 GCGTATGAAATCTCTCGCTGTCCGCTCCAGCCGCGCGGAGATGTAGCATCCCTTGG 1786
1884 GCGTATGAAATCTCTCGCTGTCCGCTCCAGCCGCGCGGAGATGTAGCATCCCTTGG 1943
1787 TCCAGCATCAGCGCTGTATAGCCCAACCCCTGTAGTGACTTAGTACCGCTTGG 1846
1944 TCCAGCATCAGCGCTGTATAGCCCAACCCCTGTAGTGACTTAGTACCGCTTGG 2003
1847 TGAATGATCAATCTATCCAGCATCAGATGCTGTGCCAAGTAGTCTCACTTGGCTT 1906
2004 TGAATGATCAATCTATCCAGCATCAGATGCTGTGCCAAGTAGTCTCACTTGGCTT 2063
1907 AAGATGTGGGCGCCAGAGAGAGAGATGATGCTGTGATTCGGATATGCCAAAGAGAG 1966
2064 AAGATGTGGGCGCCAGAGAGAGAGATGATGCTGTGATTCGGATATGCCAAAGAGAG 2123
1967 CCAGAAAGTCTCTCTTTTAAATGAAAGGTAAGTTCACACTT 2012
2124 CCAGAAAGTCTCTCTTTTAAATGAAAGGTAAGTTCACACTT 2169

RESULT 6
AAS25944
ID AAS25944 standard; cDNA; 3293 BP.
XX
AC AAS25944;
XX
DT 07-NOV-2001 (first entry)
XX
DE Human cDNA encoding a novel secreted protein, Seq ID 123.
XX
KW Human; immunosuppressive; antiarthritic; ss; antirheumatic; cytosolic;
KW cardiant; vasotropic; cerebroprotective; neurotrophic; neuroprotective;
KW antibacterial; virucide; fungicide; opthalmological; vulnerrary;
KW secreted protein; rheumatoid arthritis; hyperproliferative disorder;
KW cardiovascular disorder; cardiac arrest; cerebrovascular disorder;
KW cerebral ischaemia; angiogenesis; nervous system disorder;
KW Alzheimer's disease; infection; ocular disorder; corneal infection;
KW wound healing; epithelial cell proliferation; skin ageing; food additive;
KW preservative; antiproliferative.
XX
OS Homo sapiens.
XX
PN MO200155322-A2.
XX
PD 02-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US001341.
XX
PR 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205155P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 14-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-022513P.
PR 14-AUG-2000; 2000US-022514P.
PR 14-AUG-2000; 2000US-0225147P.
PR 14-AUG-2000; 2000US-022547P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-022681P.
PR 22-AUG-2000; 2000US-022688P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.

QY 151 CCCGAGAGACGTGATCAGTTCTTCACTTCTTAAACAATGCGCCAGGCTTGTGACGTTT 210
 DB 198 CCCGAGAGACGTGATCAGTTCTTCACTTCTTAAACAATGCGCCAGGCTTGTGACGTTT 257
 QY 211 GCGGACGTAGCCATTAACATTTTCTCAGGAGAGTGGGCGGTGTGAACCTGTGACAG 270
 DB 258 GCGGAGTGGCCATTAACATTTTCTCAGGAGAGTGGGCGGTGTGAACCTGTGACAG 317
 QY 271 GACCTGTACTGGGACGTGATGCTGAGAACTACAGTAACTTGGTCTCACTGATTTGGAG 330
 DB 318 GACCTGTACTGGGACGTGATGCTGAGAACTACAGTAACTTGGTCTCACTGATTTGGAG 377
 QY 331 TCAGCATATGAATAAATAGAGTTTACTACAGAAAAAAACATTCATGAATAAGGCTTCC 390
 DB 378 TCAGCATATGAATAAATAGAGTTTACTACAGAAAAAAACATTCATGAATAAGGCTTCC 437
 QY 391 AAAAGGAATTCAGATGAAGAAGTAATCCCTGGCGCTACAGATGTGTAAGGCTAG 450
 DB 438 AAAAGGAATTCAGATGAAGAAGTAATCCCTGGCGCTACAGATGTGTAAGGCTAG 497
 QY 451 CTTGAAAGACACAGCGCTCCAGAGGAGGTATGTCATCATGATCAATTAATGTC 510
 DB 498 CTTGAAAGACACAGCGCTCCAGAGGAGGTATGTCATCATGATCAATTAATGTC 556
 QY 511 AAAAGGCTGTCTACTAGAGAGGACCCCTCTAGAAACATCATGAGACATTAAGAG 570
 DB 557 AAAAGGCTGTCTACTAGAGAGGACCCCTCTAGAAACATCATGAGACATTAAGAG 616
 QY 571 AATTCCTTGAATGTAGAGCTGTGGGAGGCGCTTATGTGTGCTATCAACTATGCA 630
 DB 617 AATTCCTTGAATGTAGAGCTGTGGGAGGCGCTTATGTGTGCTATCAACTATGCA 676
 QY 631 CATCAGAAAAATCATACTGTGAGAAACCTTATGAATGTAAAGATGAAGAGGCTTC 690
 DB 677 CATCAGAAAAATCATACTGTGAGAAACCTTATGAATGTAAAGATGAAGAGGCTTC 736
 QY 691 CGTTGGGGCATATGACTTACTCAACATCAAAAAATTCATCTGGGGAGAAAGCCCTAC 750
 DB 737 CGTTGGGGCATATGACTTACTCAACATCAAAAAATTCATCTGGGGAGAAAGCCCTAC 796
 QY 751 TGTAAAGACGTGTGGGAGGCTTTCATGTGGGCTCAAGCGCTGTATTCATAAGAGAT 810
 DB 797 TGTAAAGACGTGTGGGAGGCTTTCATGTGGGCTCAAGCGCTGTATTCATAAGAGAT 856
 QY 811 CATACTGTGTAAAAAACCTTATGAATGTAAAGACTGTGAGAAAGGCTTTCGCGCTGTGAT 870
 DB 857 CATACTGTGTAAAAAACCTTATGAATGTAAAGACTGTGAGAAAGGCTTTCGCGCTGTGAT 915
 QY 871 GAGCTCACTGAGACCAAGAGATTCACACTGGGGAGAGAGCTACGATCCAAAGACTGT 930
 DB 916 GAGCTCACTGAGACCAAGAGATTCACACTGGGGAGAGAGCTACGATCCAAAGACTGT 975
 QY 931 GGGAGAACCTTATGCGGTGTATTAACCTATTCAGACCAAGAAATTCATATGGGGAG 990
 DB 976 GGGAGAACCTTATGCGGTGTATTAACCTATTCAGACCAAGAAATTCATATGGGGAG 1035
 QY 991 AAGCCTTACAGTGTAAAGACTGTGGGAGGCTTTTATTTGGTTCAAGCTCATTCAG 1050
 DB 1036 AAGCCTTACAGTGTAAAGACTGTGGGAGGCTTTTATTTGGTTCAAGCTCATTCAG 1095
 QY 1051 CATTAAGAAATTCACACAGGTGAGAAACCTTATGAATGTCAAGANTGTGGAGGCTTT 1110
 DB 1096 CATTAAGAAATTCACACAGGTGAGAAACCTTATGAATGTCAAGANTGTGGAGGCTTT 1155
 QY 1111 ACTCGAGTCAATTAACCTTACTCAGCANTCAAGAGTCCACCGGTGAGAGCTCCAGAA 1170
 DB 1156 ACTCGAGTCAATTAACCTTACTCAGCANTCAAGAGTCCACCGGTGAGAGCTCCAGAA 1215
 QY 1171 TGTAAAGAGTGTGGGAGGCTTTCGCTGGGCTTCAAGCTCTGTTAAGCAGAGAGATA 1230
 DB 1216 TGTAAAGAGTGTGGGAGGCTTTCGCTGGGCTTCAAGCTCTGTTAAGCAGAGAGATA 1275
 QY 1231 CATACGGGCGAGAAAGCCGTACAAAGTGCACAGAAATGTGGAGGCTTCAATGTGCTAT 1290

DB 1276 CATACGGGCGAGAAAGCCGTACAAAGTGCACAGATGTGGGAGGCTTCAATGTGCTAT 1335
 QY 1291 CACTTCACCTCAGACAGAGAAATCCACAGAGCGAAACCCCTATTAATAATGAAGAGTGT 1350
 DB 1336 CACTTCACCTCAGACAGAGAAATCCACAGAGCGAAACCCCTATTAATAATGAAGAGTGT 1395
 QY 1351 GGGAGAGCTTTCATTTATGATTCAGAGCTGTGAAAATGAGAGAAATTCATACCGGGGTG 1410
 DB 1396 GGGAGAGCTTTCATTTATGATTCAGAGCTGTGAAAATGAGAGAAATTCATACCGGGGTG 1455
 QY 1411 AAAACCTATGGGTGTACAGAAATGTGGGAGAGCTTTAGTCAAGGCCATAGCTTACAGAA 1470
 DB 1456 AAAACCTATGGGTGTACAGAAATGTGGGAGAGCTTTAGTCAAGGCCATAGCTTACAGAA 1515
 QY 1471 CATCAGAAAAACGACAGTGGGCGGAAATCTACGAATGTAAAGAGTGCAGGAGGATGT 1530
 DB 1516 CATCAGAAAAACGACAGTGGGCGGAAATCTACGAATGTAAAGAGTGCAGGAGGATGT 1575
 QY 1531 AACACCTTAACCATCTCCGAGAACATCAGAGATCCACACAGTTGAAGAGCTTTTGA 1590
 DB 1576 AACACCTTAACCATCTCCGAGAACATCAGAGATCCACACAGTTGAAGAGCTTTTGA 1635
 QY 1591 AGCAGTACCCGCTGTATCTATGTTTGGCTTCCAGATTTGTATCTGACAGTCAAC 1650
 DB 1636 AGCAGTACCCGCTGTATCTATGTTTGGCTTCCAGATTTGTATCTGACAGTCAAC 1695
 QY 1651 TGCAGTTCAAAAATATTAATGAGAAATTCAGAAAAATAAGATTTAATGTCAAAATG 1710
 DB 1696 TGCAGTTCAAAAATATTAATGAGAAATTCAGAAAAATAAGATTTAATGTCAAAATG 1755
 QY 1711 TGTGCCCTTGTAGTAGAGCTGTAGAAATCTCTGCTGTGCGGCTCCAGCCGCGGGGAT 1770
 DB 1756 TGTGCCCTTGTAGTAGAGCTGTAGAAATCTCTGCTGTGCGGCTCCAGCCGCGGGGAT 1815
 QY 1771 GTGAGTCATCCCTGTGTCCAGACATCCAGCGCTGTATAGCCACCCCTGTAGTAGAC 1830
 DB 1816 GTGAGTCATCCCTGTGTCCAGACATCCAGCGCTGTATAGCCACCCCTGTAGTAGAC 1875
 QY 1831 TTAGTAGCCGCTGTGTGTATCAGATCAATCCAGCATCAAGTGCCTGTGAGT 1890
 DB 1876 TTAGTAGCCGCTGTGTGTATCAGATCAATCCAGCATCAAGTGCCTGTGAGT 1935
 QY 1891 AGTCTCACTTGTCTTAACAGTGGCCCGAGAGACAGAGTGTGATGCTGTGATTCG 1950
 DB 1936 AGTCTCACTTGTCTTAACAGTGGCCCGAGAGACAGAGTGTGATGCTGTGATTCG 1995
 QY 1951 ATATGCCAAGAGAGGCCCAAAAGTCTCTTTTAAATGAAGAGGTGAAGTTCTCAAC 2010
 DB 1996 ATATGCCAAGAGAGGCCCAAAAGTCTCTTTTAAATGAAGAGGTGAAGTTCTCAAC 2055
 QY 2011 TTPAAA 2015
 DB 2056 TTPAAA 2060

RESULT 7
 ABX73285
 ID ABX73285 standard; DNA; 3293 BP.

ABX73285;

18-MAR-2003 (first entry)

Human novel polynucleotide #113.

DE Human; gene; de; neural disorder; immune system disorder; renal disorder;
 XX muscular disorder; respiratory disease; reproductive disorder;
 XX gastroenteric disorder; pulmonary disorder; cardiovascular disorder;
 KM hyperproliferative disorder; inflammatory disease; allergic reaction;
 KM blood related disorder; cancer; immunosuppressive; anti-inflammatory;
 KM cardiovascular; nephrotoxic; cytostatic; antiallergic; thrombolytic;
 KM haemostatic; antiarteriosclerotic.

XX OS Homo sapiens.
XX PN US2002132753-A1.
XX PD 19-SEP-2002.
XX PE 17-JAN-2001; 2001US-00764864.
XX 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 28-JUN-2000; 2000US-0214865P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 22-AUG-2000; 2000US-0226868P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 08-SEP-2000; 2000US-0231413P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234977P.
PR 27-SEP-2000; 2000US-0235834P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239355P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241809P.
PR 01-NOV-2000; 2000US-0244617P.
PR 17-NOV-2000; 2000US-0249299P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
XX (ROSE/) ROSEN C A.
PA (RUBE/) RUBEN S M.
PA (BARA/) BARASH S C.
XX
PI Rosen CA, Ruben SM, Barash SC,
XX MPI: 2003-147444/14.
DR P-PSDB; ABUS5025.
XX
PT New polypeptides and nucleic acids, useful in gene therapy for treating,
PT inhibiting or preventing e.g. neural, immune system, muscular,
PT respiratory, reproductive, gastrointestinal, pulmonary, cardiovascular or
PT renal disorders.

XX PS Claim 1; SEQ ID NO 123; 402bp; English.
XX CC The invention relates to human novel polypeptides and their associated
CC polynucleotides. The polypeptides and polynucleotides are useful in gene
CC therapy for treating, inhibiting or preventing neural disorders, immune
CC system disorders (e.g. systemic lupus erythematosus, rheumatoid arthritis
CC and multiple sclerosis), muscular disorders, respiratory diseases (e.g.
CC nasal vestibulitis, nasal polyps and sinusitis), reproductive disorders,
CC gastrointestinal disorders, pulmonary disorders, cardiovascular disorders
CC (e.g. congenital heart defects, Ebstein's anomaly and hypoplastic left
CC heart syndrome), renal disorders (e.g. acute kidney failure and end-stage
CC renal disease), hyperproliferative disorders (e.g. Hodgkin's disease and
CC leukaemia), inflammatory diseases (e.g. septic shock, bursitis and
CC appendicitis), allergic reactions and conditions (e.g. asthma), blood
CC related disorders (e.g. thrombosis, atherosclerosis and myocardial
CC infarction) and cancerous diseases. Sequences ABX73173-ABX74167 represent
XX human novel polynucleotides of the invention

Sequence 3293 BP; 974 A; 717 C; 750 G; 852 T; 0 U; 0 Other;
Query Match 96.6%; Score 1946.2; DB 8; Length 3293;
Best Local Similarity 98.3%; Pred. No. 0;
Matches 2010; Conservative 0; Mismatches 3; Indels 32; Gaps 3;

QY 1 CGCTTTCACCCGTCAGCAACCGCTGTCTCCGCCATCCCGGCGCAGCATCTCT 60
DB CGCTTTCGACCCGTCAGCAACCGCTGTCTCTCCGCCATCCCGGCGCAGCATCTCT 77
QY 61 TCAGAAAAGCATCCCGAGGAGGAAGCAATGGTTAAATCATCTAGTCTC----- 111
DB TCAGAAAAGCATCCCGAGGAGGAAGCAATGGTTAAATCATCTAGTCTC----- 137
QY 78 TCAGAAAAGCATCCCGAGGAGGAAGCAATGGTTAAATCATCTAGTCTC----- 151
DB TCAGAAAAGCATCCCGAGGAGGAAGCAATGGTTAAATCATCTAGTCTC----- 197
QY 112 -----AGCTTACCTCTCGAATTTGTTCTTTCAGTGAAGAAC 150
DB AGTGAGAAACCGACACTGTGAGGCTTACCTTCGGAATTTGTTCTTTCAGTGAAGAAC 197
QY 138 AGTGAGAAACCGACACTGTGAGGCTTACCTTCGGAATTTGTTCTTTCAGTGAAGAAC 197
DB AGTGAGAAACCGACACTGTGAGGCTTACCTTCGGAATTTGTTCTTTCAGTGAAGAAC 257
QY 151 CCCGAGAAAGTCACTGATCACTTTCTTCAAGTCTTAAACATAGCCCGGCTTGGTACCTTC 210
DB CCCGAGAAAGTCACTGATCACTTTCTTCAAGTCTTAAACATAGCCCGGCTTGGTACCTTC 257
QY 198 CCCGAGAAAGTCACTGATCACTTTCTTCAAGTCTTAAACATAGCCCGGCTTGGTACCTTC 257
DB CCCGAGAAAGTCACTGATCACTTTCTTCAAGTCTTAAACATAGCCCGGCTTGGTACCTTC 270
QY 211 GCCGAGTACGATCACTGATCACTTTCTTCAAGTCTTAAACATAGCCCGGCTTGGTACCTTC 270
DB GCCGAGTACGATCACTGATCACTTTCTTCAAGTCTTAAACATAGCCCGGCTTGGTACCTTC 317
QY 258 GCCGAGTACGATCACTGATCACTTTCTTCAAGTCTTAAACATAGCCCGGCTTGGTACCTTC 317
DB GCCGAGTACGATCACTGATCACTTTCTTCAAGTCTTAAACATAGCCCGGCTTGGTACCTTC 330
QY 271 GACCTGTACTGGAGCGTGATCTGGAGAACTTACAGTAACTTGGTCTCACTGGATTGGAG 330
DB GACCTGTACTGGAGCGTGATCTGGAGAACTTACAGTAACTTGGTCTCACTGGATTGGAG 377
QY 318 GACCTGTACTGGAGCGTGATCTGGAGAACTTACAGTAACTTGGTCTCACTGGATTGGAG 377
DB GACCTGTACTGGAGCGTGATCTGGAGAACTTACAGTAACTTGGTCTCACTGGATTGGAG 390
QY 331 TCAGCATATGAAAATTAAGAGTTTACCTTACAGAAAATTAAGAGTTTACCTTACAGAAA 390
DB TCAGCATATGAAAATTAAGAGTTTACCTTACAGAAAATTAAGAGTTTACCTTACAGAAA 437
QY 378 TCAGCATATGAAAATTAAGAGTTTACCTTACAGAAAATTAAGAGTTTACCTTACAGAAA 437
DB TCAGCATATGAAAATTAAGAGTTTACCTTACAGAAAATTAAGAGTTTACCTTACAGAAA 450
QY 391 AAAAGAAATTCAGATAGAGAAAGTAAATCCCTTGGCGGTAATCGGATATGGAAGTACG 450
DB AAAAGAAATTCAGATAGAGAAAGTAAATCCCTTGGCGGTAATCGGATATGGAAGTACG 497
QY 438 AAAAGAAATTCAGATAGAGAAAGTAAATCCCTTGGCGGTAATCGGATATGGAAGTACG 497
DB AAAAGAAATTCAGATAGAGAAAGTAAATCCCTTGGCGGTAATCGGATATGGAAGTACG 510
QY 451 CTTGAAAGACCAAGCGCTCCAGAGGAGTATGTCATCATGATCATCAATTAATGTC 510
DB CTTGAAAGACCAAGCGCTCCAGAGGAGTATGTCATCATGATCATCAATTAATGTC 556
QY 498 CTTGAAAGACCAAGCGCTCCAGAGGAGTATGTCATCATGATCATCAATTAATGTC 556
DB CTTGAAAGACCAAGCGCTCCAGAGGAGTATGTCATCATGATCATCAATTAATGTC 570
QY 511 AAAAGCGCTGCTTACTAGAGAGGACCCCTTCTTGAACAATCATGAGACATCATTAAGAG 570
DB AAAAGCGCTGCTTACTAGAGAGGACCCCTTCTTGAACAATCATGAGACATCATTAAGAG 616
QY 557 AAAAGCGCTGCTTACTAGAGAGGACCCCTTCTTGAACAATCATGAGACATCATTAAGAG 616
DB AAAAGCGCTGCTTACTAGAGAGGACCCCTTCTTGAACAATCATGAGACATCATTAAGAG 630
QY 571 AATTCCTTGAATGTAAGAGTGTGGAGAGGCTTATGTCGAGTCACTTAAGTCA 630
DB AATTCCTTGAATGTAAGAGTGTGGAGAGGCTTATGTCGAGTCACTTAAGTCA 676
QY 617 AATTCCTTGAATGTAAGAGTGTGGAGAGGCTTATGTCGAGTCACTTAAGTCA 676
DB AATTCCTTGAATGTAAGAGTGTGGAGAGGCTTATGTCGAGTCACTTAAGTCA 690
QY 631 CATCAGAAATTCATATCTGTGAGAAACCTTATGAAATGTAAGATGTAAGAGGCTTTC 690
DB CATCAGAAATTCATATCTGTGAGAAACCTTATGAAATGTAAGATGTAAGAGGCTTTC 736
QY 677 CATCAGAAATTCATATCTGTGAGAAACCTTATGAAATGTAAGATGTAAGAGGCTTTC 736
DB CATCAGAAATTCATATCTGTGAGAAACCTTATGAAATGTAAGATGTAAGAGGCTTTC 766

```

QY 691 CGTGGGGCAATCAGCTTACTCAACATCAAAAAATTCATCTGGGGAGAGCCCTACGAA 750
    |||||
Db 737 CATTGGGGCAATCAGCTTACTCAACATCAAAAAATTCATCTGGGGAGAGCCCTACGAA 796
    |||||
QY 751 TGTAAAGACTGTGGAGAGGCTTTTGAATGGGGCTCAAGCTCGTTATTTCAATAAGAGATT 810
    |||||
Db 797 TGTAAAGACTGTGGAGAGGCTTTTGAATGGGGCTCAAGCTCGTTATTTCAATAAGAGATT 856
    |||||
QY 811 CATATCGTGAAGAAACCTATGATGTAAAGACTGTGAAGAGGCTTTGGCGTGTGAT 870
    |||||
Db 857 CATATCGTGAAGAAACCTATGATGTAAAGACTGTGAAGAGGCTTTGGCGTGTGAT 915
    |||||
QY 871 GAGCTCACTCAGCAGCAGAGATTCCA CACTGGGAGAGAAAGACTCAAGATGCAGAACTGT 930
    |||||
Db 916 GAGCTCACTCAGCAGCAGAGATTCCA CACTGGGAGAGAAAGACTCAAGATGCAGAACTGT 975
    |||||
QY 931 GGGAGAGCCTTTAGCCGTGTGTATTAACCTTATCAGACAGAGAAATTCATAGTGGGGAG 990
    |||||
Db 976 GGGAGAGCCTTTAGCCGTGTGTATTAACCTTATCAGACAGAGAAATTCATAGTGGGGAG 1035
    |||||
QY 991 AAGCCTTACGAGTGTAAAGACTGTGGAGAGGCTTTTATTTGTGTTCAAGCTCATTTGAG 1050
    |||||
Db 1036 AAGCCTTACGAGTGTAAAGACTGTGGAGAGGCTTTTATTTGTGTTCAAGCTCATTTGAG 1095
    |||||
QY 1051 CATAAAGAAATTCACACAGGTGAGAAACCTATGAAATGTCAAGAAATGTGGAGAGGCTTT 1110
    |||||
Db 1096 CATAAAGAAATTCACACAGGTGAGAAACCTATGAAATGTCAAGAAATGTGGAGAGGCTTT 1155
    |||||
QY 1111 ACTCGAGTCAATTTACTACTGAGCATCAAGAAATCCACACCGGTGAGAGAGCTTCACGAA 1170
    |||||
Db 1156 ACTCGAGTCAATTTACTACTGAGCATCAAGAAATCCACACCGGTGAGAGAGCTTCACGAA 1215
    |||||
QY 1171 TGTAAAGAGTGTGGAGAGGCTTTTGGTGGGTTTCGAGCTCGTTAAGCAGAGAGAAATTA 1230
    |||||
Db 1216 TGTAAAGAGTGTGGAGAGGCTTTTGGTGGGTTTCGAGCTCGTTAAGCAGAGAGAAATTA 1275
    |||||
QY 1231 CATACGGGCGAGAGCCGTACAAGTGCACAGAAATGTGGAGAGGCTTCAATGTGCTAT 1290
    |||||
Db 1276 CATACGGGCGAGAGCCGTACAAGTGCACAGAAATGTGGAGAGGCTTCAATGTGCTAT 1335
    |||||
QY 1291 CACCTCACTCAGACGAGAGAAATCCACAGAGCGGAAACCCCGTAAATGTAGAGAGTGT 1350
    |||||
Db 1336 CACCTCACTCAGACGAGAGAAATCCACAGAGCGGAAACCCCGTAAATGTAGAGAGTGT 1395
    |||||
QY 1351 GGGAGAGGCTTTTCAATTTATGAGTCAAGCTCGTGAACAATGAGAGAAATTCACCGGGGTTG 1410
    |||||
Db 1396 GGGAGAGGCTTTTCAATTTATGAGTCAAGCTCGTGAACAATGAGAGAAATTCACCGGGGTTG 1455
    |||||
QY 1411 AAACCTATGGTGTACAGAAATGTGGAGAGGCTTTTGTACAGCGCCATCAGCTTACACAA 1470
    |||||
Db 1456 AAACCTATGGTGTGTACAGAAATGTGGAGAGGCTTTTGTACAGCGCCATCAGCTTACACAA 1515
    |||||
QY 1471 CATCAGAAAAACGACAGTGGGGGAGAAATCTTACGAATGTAAAGAGTGGGGAGGAGTGT 1530
    |||||
Db 1516 CATCAGAAAAACGACAGTGGGGGAGAAATCTTACGAATGTAAAGAGTGGGGAGGAGTGT 1575
    |||||
QY 1531 AACCACTAAACCATCTCCGAGAACATCAGAGATCCACAAAGTTGAAGAGCTTTTGA 1590
    |||||
Db 1576 AACCACTAAACCATCTCCGAGAACATCAGAGATCCACAAAGTTGAAGAGCTTTTGA 1635
    |||||
QY 1591 ACGCAGTAGCCGCTCGTATCTATGTGTTTCCAGATTGTTTACTGTGAGTCAAC 1650
    |||||
Db 1636 ACGCAGTAGCCGCTCGTATCTATGTGTTTCCAGATTGTTTACTGTGAGTCAAC 1695
    |||||
QY 1651 TGCAGTTCAAAAATTTAATGTGAATAATTCAGAAAAAAGAAATTTTAATCTCAATGG 1710
    |||||
Db 1696 TGCAGTTCAAAAATTTAATGTGAATAATTCAGAAAAAAGAAATTTTAATCTCAATGG 1755
    |||||
QY 1711 TGTGCGCTTGTGAGTAGCTGATGAATCTCGCTGTCCGGCTCCAGCGCGCGGGGAGT 1770
    |||||
Db 1756 TGTGCGCTTGTGAGTAGCTGATGAATCTCGCTGTCCGGCTCCAGCGCGCGGGGAGT 1815
    |||||
QY 1771 GTGAGTCAATCCCTGTGTCCAGACATCAGCGTGTATAGCCACCCCTGTGTAAGTAC 1830
    |||||

```

```

Db 1816 GTGAGTCAATCCCTGTGTCCAGACATCCAGCTGTATAGCCACCCACTGTCTAGTAC 1875
    |||||
QY 1831 TTAGTAGCCGCTCTTGATGATCAGATCAACTATCCAGCATCAGAGTCTGTGCCAAGT 1890
    |||||
Db 1876 TTAGTAGCCGCTCTTGATGATCAGATCAACTATCCAGCATCAGAGTCTGTGCCAAGT 1935
    |||||
QY 1891 AGTCTCACTTTGCTTAAACAGTGGCCCAAGAGAGAGAGTGAATGTGTGATTCGG 1950
    |||||
Db 1936 AGTCTCACTTTGCTTAAACAGTGGCCCAAGAGAGAGAGTGAATGTGTGATTCGG 1995
    |||||
QY 1951 ATATGCCAAAGAGAGCCCAAGTGTCTCTTTTAAATGAAAAGTTCAAC 2010
    |||||
Db 1996 ATATGCCAAAGAGAGCCCAAGTGTCTCTTTTAAATGAAAAGTTCAAC 2055
    |||||
QY 2011 TTTAA 2015
    |||||
Db 2056 TTTAA 2060
    |||||

RESULT 8
ABNS9917
ID ABNS9917 standard; cDNA; 2010 BP.
XX
AC ABNS9917;
XX
DT 28-JUN-2002 (first entry)
XX
DE Novel human coding sequence SEQ ID NO: 328.
XX
KW Human; anti-naemic; vulnerrary; anti-inflammatory; immunomodulator;
KW anti-infectility; cerebroprotective; cytostatic; rheumatic; gene therapy;
KW neuroprotective; antiparkinsonian; protein therapy; EST;
KW expressed sequence tag; gene; ss.
XX
OS Homo sapiens.
XX
PN W0200222660-A2.
XX
PD 21-MAR-2002.
XX
PF 10-SEP-2001; 2001WO-US026015.
XX
PR 11-SEP-2000; 2000US-00659671.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Liu C, Zhou P, Asundi V, Zhang J, Zhao QX, Ren F;
PI Xue AJ, Yang Y, Wenman T, Drmanac RT;
XX
DR WPI; 2002-292408/33.
XX
DR P-PSDB; ABB97504.
XX
PT An isolated polynucleotide for treating diseases associated with its
XX encoded polypeptide such as cancer and multiple sclerosis.
XX
PS Claim 1; SEQ ID NO 328; 509pp; English.
XX
CC The present invention provides the protein and coding sequences of 444
CC novel human proteins. These were isolated from expressed sequences tags
CC (ESTs). They can be used to stimulate cell growth, to regulate
CC haematopoiesis e.g. to treat aplastic anemia, to help tissue regrowth
CC e.g. in burn treatment, to regulate the immune system e.g. to treat
CC multiple sclerosis, to regulate activin or inhibin e.g. to treat
CC infertility, to regulate haemostasis or thrombolysis e.g. to treat stroke
CC and cancer, to screen for drugs, to treat inflammatory conditions e.g.
CC rheumatoid arthritis, and to treat nervous system disorders e.g.
CC Parkinson's disease. The present sequence is a coding sequence of the
CC invention
XX
SQ Sequence 2010 BP; 585 A; 446 C; 527 G; 452 T; 0 U; 0 Other;

```

Query Match 87.7%; Score 1768; DB 6; Length 2010;

Best Local Similarity 100.0%; Pred. No. 0;
Matches 1768; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 47 CAGGCCAGCATCTCTTACAGAAAAAGCATCCCGAGAGAGAAAGCAATCGTTAAACATCTT 106
Db 243 CAGGCCAGCATCTCTTACAGAAAAAGCATCCCGAGAGAGAAAGCAATCGTTAAACATCTT 102
QY 107 AGGTGAGCTCTAGCCCTCTCGGAAATTTGCTCTTCACTGAGAAACCCCGAAGAACTGATC 166
Db 303 AGGTGAGCTCTAGCCCTCTCGGAAATTTGCTCTTCACTGAGAAACCCCGAAGAACTGATC 362
QY 167 AGTTCTTCAGTTCTTAAACAAATGAGCCAGAGTTTGGTGAAGTTCCGCGAAGTACAG 226
Db 363 AGTTCTTCAGTTCTTAAACAAATGAGCCAGAGTTTGGTGAAGTTCCGCGAAGTACAG 422
QY 227 ACTTTTCTCAGAGAGAGTGGGCTCTGTCTGAACCTCTGCTCAGAGAGACCTGTACTGG 286
Db 423 ACTTTTCTCAGAGAGAGTGGGCTCTGTCTGAACCTCTGCTCAGAGAGACCTGTACTGG 482
QY 287 TGATGCTGAGAACTACAGTAAGTGTCTCACTGATTTGGAGTACAGATATGAAAAA 346
Db 483 TGATGCTGAGAACTACAGTAAGTGTCTCACTGATTTGGAGTACAGATATGAAAAA 542
QY 347 AGAGTTTACCTACAGAAAAAAACATTCATGAATAAGGCTTCCAAAAGAAATTCAGATA 406
Db 543 AGAGTTTACCTACAGAAAAAAACATTCATGAATAAGGCTTCCAAAAGAAATTCAGATA 602
QY 407 GAAAGATTAATTCCTTGGCCGTAACTGGAATATGTGAAGATAAGCTTGAAAGACCAAGC 466
Db 603 GAAAGATTAATTCCTTGGCCGTAACTGGAATATGTGAAGATAAGCTTGAAAGACCAAGC 662
QY 467 GCTCCAGAGGAGGATATGTCATTCAGATGATCATCAATATGTCAAAAGGCTCGTACTA 526
Db 663 GCTCCAGAGGAGGATATGTCATTCAGATGATCATCAATATGTCAAAAGGCTCGTACTA 722
QY 527 GAGAAGGCAACCCCTCTAGAAACATCATCAGACATCATTAAGAGAAATTCCTTGAATGA 586
Db 723 GAGAAGGCAACCCCTCTAGAAACATCATCAGACATCATTAAGAGAAATTCCTTGAATGA 782
QY 587 AGAGCTGTGGGAAGGCTTTAGTCTGTGCTATCAACTTAAGTCAACATCAGAAATCCATA 646
Db 783 AGAGCTGTGGGAAGGCTTTAGTCTGTGCTATCAACTTAAGTCAACATCAGAAATCCATA 842
QY 647 CTGTGAGAGAAACCTTATGAAATGTAAGAAATGTAAGAGGCTTCCGTTGGGCAATCAGC 706
Db 843 CTGTGAGAGAAACCTTATGAAATGTAAGAAATGTAAGAGGCTTCCGTTGGGCAATCAGC 902
QY 707 TTAACCAACATCAAAAAATTCATACCTGGGAGAGAGCCCTACGAATGTAAAGACTGTGGGA 766
Db 903 TTAACCAACATCAAAAAATTCATACCTGGGAGAGAGCCCTACGAATGTAAAGACTGTGGGA 962
QY 767 AGGCTTTTCGATGGGCTCAAGCCCTGTTATTCATTAAGAGATTCATATCTGTGAAAAAC 826
Db 963 AGGCTTTTCGATGGGCTCAAGCCCTGTTATTCATTAAGAGATTCATATCTGTGAAAAAC 1022
QY 827 CCTATGAATGTAAAGACTGTGAAAGGCTTTCGGCGTGTGATGAGCTCATCAGCAC 886
Db 1023 CCTATGAATGTAAAGACTGTGAAAGGCTTTCGGCGTGTGATGAGCTCATCAGCAC 1082
QY 887 AGAGATTCACACCTGGGAGAGAAAGTACGAATGCAAGACTGTGGGAAAGCCTTTAGCC 946
Db 1083 AGAGATTCACACCTGGGAGAGAAAGTACGAATGCAAGACTGTGGGAAAGCCTTTAGCC 1142
QY 947 GTGTGTATTAACCTTATTCAGCAAGAGAAATTCATAGTGGGAGAGACCTTTCAGAGTGA 1006
Db 1143 GTGTGTATTAACCTTATTCAGCAAGAGAAATTCATAGTGGGAGAGACCTTTCAGAGTGA 1202
QY 1007 AAGACTGTGGGAAGGCTTTATTTGGTTCAGGCTCATTTAGAGATTAAGAAATTCACA 1066
Db 1203 AAGACTGTGGGAAGGCTTTATTTGGTTCAGGCTCATTTAGAGATTAAGAAATTCACA 1262
QY 1067 CAGGTGAGAAACCTTATGATGTCAAGATGTGGAGAGGCTTTTACTCGAGTCAATTACC 1126

```

```

Db 1263 CAGGTGAGAAACCTTATGATGTCAAGATGTGGAGAGGCTTTTACTCGAGTCAATTACC 1322
QY 1127 TTAACAGCATCAGAAAGATCCACACCGGTGAGAGCTTCACGAATGTAAAGAGTGTGGGA 1186
Db 1323 TTAACAGCATCAGAAAGATCCACACCGGTGAGAGCTTCACGAATGTAAAGAGTGTGGGA 1382
QY 1187 AGGCTTTTCGCGGGGTTTCAGAGCTCGTTAAGCAGAGAGATACATACGGGCGAAGAC 1246
Db 1383 AGGCTTTTCGCGGGGTTTCAGAGCTCGTTAAGCAGAGAGATACATACGGGCGAAGAC 1442
QY 1247 CGTCAAGTGCACAGAAATGTGGGAAGGCTTCAATTTGTGGCTATCACTCACTCAGCAG 1306
Db 1443 CGTCAAGTGCACAGAAATGTGGGAAGGCTTCAATTTGTGGCTATCACTCACTCAGCAG 1502
QY 1307 AGAGATTCACACAGGCGGAAACCCCGTATTAATGTAAAGAGTGTGGAAAGGCTTTCATT 1366
Db 1503 AGAGATTCACACAGGCGGAAACCCCGTATTAATGTAAAGAGTGTGGAAAGGCTTTCATT 1562
QY 1367 ATGATCGAGCCTCGTGAACATGAGAGATTCATACCGGGGGTGAACCTTATGGGTGA 1426
Db 1563 ATGATCGAGCCTCGTGAACATGAGAGATTCATACCGGGGGTGAACCTTATGGGTGA 1622
QY 1427 CAGAAATGTGGGAAGAGCTTTAGTCAAGGCTCATCACTTACACAGATCAGAAAAAGCACA 1486
Db 1623 CAGAAATGTGGGAAGAGCTTTAGTCAAGGCTCATCACTTACACAGATCAGAAAAAGCACA 1682
QY 1487 GTGGGGCGAAATCCTACGAATGTAAAGAGTGGGGAAGGCAATGTAAACCTTAAACATC 1546
Db 1683 GTGGGGCGAAATCCTACGAATGTAAAGAGTGGGGAAGGCAATGTAAACCTTAAACATC 1742
QY 1547 TCCGGAACATCAGAGAGATCCACACAGTGAAGAGCCTTTTGAACGAGTGGCCGCTC 1606
Db 1743 TCCGGAACATCAGAGAGATCCACACAGTGAAGAGCCTTTTGAACGAGTGGCCGCTC 1802
QY 1607 GTATCTATGTTTCCCTTCCACAGTTTGTATCTGCACTCACTGCACTGCACTTCAAAAAAT 1666
Db 1803 GTATCTATGTTTCCCTTCCACAGTTTGTATCTGCACTCACTGCACTGCACTTCAAAAAAT 1862
QY 1667 TAAATGAAAAATTCAGAAATTAAGAAATTTTAAGTCTCAATGTGTGCTTCTGAGTA 1726
Db 1863 TAAATGAAAAATTCAGAAATTAAGAAATTTTAAGTCTCAATGTGTGCTTCTGAGTA 1922
QY 1727 GCGGTATGAATTCCTCGCTGCGGCTCCAGCCGCGGGGAGATGAGTCAATCCCTTGG 1786
Db 1923 GCGGTATGAATTCCTCGCTGCGGCTCCAGCCGCGGGGAGATGAGTCAATCCCTTGG 1982
QY 1787 TCCAGCACATCCACGCTGTATACGCCAC 1814
Db 1983 TCCAGCACATCCACGCTGTATACGCCAC 2010

```

RESULT 9
AAH9589/c
ID AAH9589 standard; cDNA, 1453 BP.

AAH9589;

16-OCT-2001 (first entry)

Human protein encoding cDNA sequence SEQ ID NO:424.

Human, cancer; HIV infection; human immunodeficiency virus;
antibacterial; antineoplastic; antiarthritic; immunosuppressive;
anti-infective; endocrine; cardiac; central nervous system; virucide;
anti-HIV; fungicide; antimutagen; cardiovascular; antineoplastic; anaemia;
antiallergic; haemostatic; antileukemic; antileukemic; osteoporosis; eczema;
dermatological; antiallergic; antineoplastic; antidiabetic; cytotoxic;
neuroprotective; antidepressant; neurotrophic; antiparkinsonian; infection;
immunostimulant; gene therapy; antisense therapy; vaccine; inflammation;
anti-anaphylactic; rheumatoid arthritis; septic shock; pancreatitis;
cardiac dysfunction; neuropathology; cardiac anaphylaxis; autoimmunity;
genetic disease; haematopoietic disorder; platelet disorder; asthma;
thrombocytopenia; osteoporosis; severe combined immunodeficiency;

KW allergic rhinitis; diabetes; multiple sclerosis; depression;
 KM Alzheimer's disease; Parkinson's disease; neurodegenerative disorder;
 KM neurological disorder; ss.
 OS Homo sapiens.
 PN W0200153455-A2.
 XX 26-JUL-2001.
 PD 22-DEC-2000; 2000MO-US035017.
 PF 23-DEC-1999; 99US-00471275.
 XX 21-JAN-2000; 2000US-00488725.
 PR 25-APR-2000; 2000US-00552317.
 XX (HUSE-) HUSEQ INC.
 PA Tang YT, Liu C, Drmanac RT;
 PI WPI: 2001-457603/49.
 XX P-PSDB; AAM25648.
 DR Isolated human polynucleotides encoding polypeptides, useful for the
 PT treatment and diagnosis of e.g. cancer, ulcers and HIV infection.
 XX Claim 1; Page 516-517; 1217pp; English.
 PS AAM99166 to AAM99904 encode the human proteins given in AAM2525 to
 CC AAM25966. The proteins can have activities based on the tissues and cells
 CC they are expressed in, such as: antiinflammatory; antineumatic;
 CC antiarthritic; immunosuppressive; antibacterial; endocrine; cardiac;
 CC central nervous system; virucide; anti-HIV; fungicide; antimutagen;
 CC cardiovascular; antianemic; antiaggregant; haemostatic; vulnerary;
 CC antitumor; osteopathic; dermatological; antiallergic; antiautomatic;
 CC antidiabetic; cytostatic; neuroprotective; antidepressant; nootropic;
 CC antiparkinsonian; and immunostimulant. The proteins and polynucleotides
 CC encoding them can be used in gene therapy, antisense therapy and vaccine
 CC production. The proteins and polynucleotides are useful for screening for
 CC agonists or antagonists of a protein and for the treatment and diagnosis
 CC of disorders associated with the activity of a protein e.g. inflammation,
 CC rheumatoid arthritis, septic shock, pancreatitis, cardiac dysfunction,
 CC neuropathology, cardiac anaphylaxis, viral, bacterial, HIV and fungal
 CC infections, autoimmunity, genetic diseases, haematopoietic disorders,
 CC anaemia, platelet disorders, thrombocytopenia, wounds, burns, ulcers,
 CC osteoporosis, severe combined immunodeficiency, eczema, allergic
 CC rhinitis, asthma, diabetes, cancer, multiple sclerosis, depression,
 CC Alzheimer's disease, Parkinson's disease, neurodegenerative and
 CC neurological disorders
 XX Sequence 1453 BP; 340 A; 366 C; 312 G; 435 T; 0 U; 0 Other;
 SQ
 Query Match 70.5%; Score 1420.2; DB 4; Length 1453;
 Best Local Similarity 99.5%; Pred. No. 0;
 Matches 1446; Conservative 0; Mismatches 3; Indels 4; Gaps 2;

Db 1213 GATTCACTAGTGGTGAACAAACCTTATGAATGTAAAGACTGTGAAAGGCTTTTCGCGTGG 1154
 |||
 QY TGAATAGCTCACTGACGACCAAGATTCACACATCTGGGAGAAAGACTACGAATGCAAGA 926
 |||
 Db 1153 TGATAGAGCTCACTGACGACCAAGATTCACACATCTGGGAGAAAGACTACGAATGCAAGA 1094
 |||
 QY CTGTGGGAAGACCTTTAGCCGTGTGTATTAATTATTCAGCAACAAGAAATTCATAGTGG 986
 |||
 Db 1093 CTGTGGGAAGACCTTTAGCCGTGTGTATTAATTATTCAGCAACAAGAAATTCATAGTGG 1034
 |||
 QY GGAGAAAGCTTTACGAGTGTAAAGACTGTGGAAAGCTTTATTTGTGTTCAAGCTCAT 1046
 |||
 Db 1033 GGAGAAAGCTTTACGAGTGTAAAGACTGTGGAAAGCTTTATTTGTGTTCAAGCTCAT 974
 |||
 QY 1047 TCAGCATTAAGAAATTCACAGGTGTGAAGAAACCTATGAATGTCAAGAAATGTGGAGGC 1106
 |||
 Db 973 TCAGCATTAAGAAATTCACAGGTGTGAAGAAACCTATGAATGTCAAGAAATGTGGAGGC 914
 |||
 QY 1107 CTTTACTCGAGTCAATTAACCTTACTCAGCATGAAATCCACACCGGTGAAGAGCTTCA 1166
 |||
 Db 913 CTTTACTCGAGTCAATTAACCTTACTCAGCATGAAATCCACACCGGTGAAGAGCTTCA 854
 |||
 QY 1167 CGAATGTAAAGAGTGTGGAAAGGCTTTTGTGGGGTTTGAGCTTCGTTTAAGCAGAGG 1226
 |||
 Db 853 CGAATGTAAAGAGTGTGGAAAGGCTTTTGTGGGGTTTGAGCTTCGTTTAAGCAGAGG 794
 |||
 QY 1227 GATPACATACGGGCGAAGAGCGGTACAAAGTGCACAGAAATGTGGAAAGGCTTCAATTGTGG 1286
 |||
 Db 793 GATPACATACGGGCGAAGAGCGGTACAAAGTGCACAGAAATGTGGAAAGGCTTCAATTGTGG 734
 |||
 QY 1287 CTATCACTCACTGACGACGAGAGATCCACACAGGCGAAACCCCGTATTAATGTAAAGGA 1346
 |||
 Db 733 CTATCACTCACTGACGACGAGAGATCCACACAGGCGAAACCCCGTATTAATGTAAAGGA 674
 |||
 QY 1347 GTGTGGAAAGCTTTCAATTATGTATCGAGCTCGTGAACATGAGAAATTCATACCGG 1406
 |||
 Db 673 GTGTGGAAAGCTTTCAATTATGTATCGAGCTCGTGAACATGAGAAATTCATACCGG 614
 |||
 QY 1407 GGTTGAACCCCTTAATGGGTGTACAGAAATGTGGAAAGGCTTTAGTCACGGCATCAGCTTAC 1466
 |||
 Db 613 GGTTGAACCCCTTAATGGGTGTACAGAAATGTGGAAAGGCTTTAGTCACGGCATCAGCTTAC 554
 |||
 QY 1467 ACNACATCGAAGAAACGCAAGTGGGCGAAATCTTACGATGTAAAGAGTGGCGGAAGGC 1526
 |||
 Db 553 ACNACATCGAAGAAACGCAAGTGGGCGAAATCTTACGATGTAAAGAGTGGCGGAAGGC 494
 |||
 QY 1527 ATGTAAACAATTAAACCATCTCCGAGACATCAGAGATCCACACAGTTGAAGAGCTTT 1586
 |||
 Db 493 ATGTAAACAATTAAACCATCTCCGAGACATCAGAGATCCACACAGTTGAAGAGCTTT 434
 |||
 QY 1587 TTGAAGCGAGTGGCCGCGTCTGATCTATGTTTGGCTTCCACAGTTTGTATCTGACAGT 1646
 |||
 Db 433 TTGAAGCGAGTGGCCGCGTCTGATCTATGTTTGGCTTCCACAGTTTGTATCTGACAGT 374
 |||
 QY 1647 CAATCGAGTTCAAAATATTAATGAAGAAATTCAGAAATTAAGAAATTTAAGTCTCAA 1706
 |||
 Db 373 CAATCGAGTTCAAAATATTAATGAAGAAATTCAGAAATTAAGAAATTTAAGTCTCAA 314
 |||
 QY 1707 ATGTGTGCCCTTTCTGAGTACGTGATGAATCTTCGCTGTCCGCTCCAGCCGCGCG 1766
 |||
 Db 313 ATGTGTGCCCTTTCTGAGTACGTGATGAATCTTCGCTGTCCGCTCCAGCCGCGCG 254
 |||
 QY 1767 GGAATGAGTCAATCCCTGGTCCAGACATCCAGCGTATATACGCAACCAACCTGCTAG 1826
 |||
 Db 253 GGAATGAGTCAATCCCTGGTCCAGACATCCAGCGTATATACGCAACCAACCTGCTAG 194
 |||
 QY 1827 TGACTTAGTACCGCTCTTGATGATCAGATCAATATCCAGCA--TCAAGTGTGCTGTGCC 1885
 |||
 Db 193 TGACTTAGTACCGCTCTTGATGATCAGATCAATATCCAGCA--TCAAGTGTGCTGTGAC 134
 |||
 QY 1886 CAAATAGTCTCACTTTGCTTAAACAGTGGCCCAAGAGAGAGAGTNG--TGAATGCTGG 1942
 |||

Db 133 CAAGCAGTCCCTGCTTGGCTTAACAGTGGCCCCAGAGACAGAGTGAATGATGCTGG 74
QY 1943 TGATTCGATATGCCAAGAGAGAGCCACAAAGTCTCTTTTAATGAAAAGGTGAAG 2002
Db 73 TGATTCGATATGCCAAGAGAGAGCCACAAAGTCTCTTTTAATGAAAAGGTGAAG 14
QY 2003 TTCTCACTTAA 2015
Db 13 TTCTCACTTAA 1
RESULT 10
AD13363
ID AD13363 standard; cDNA; 1429 BP.
XX AD13363;
AC AD13363;
XX 17-JUN-2004 (first entry)
XX 17-JUN-2004 (first entry)
DE Human cDNA #689.
XX Human; gene; ss; immunological response; immunopathological condition;
KM Crohn's disease; asthma; ulcerative colitis; hypereosinophilia;
KM irritable bowel syndrome; osteoarthritis; rheumatoid arthritis;
KM acute monocytic leukemia; antiinflammatory; antiasthmatic; anti-leuk;
KM osteopathic; antiarthritic; antirheumatic; cytostatic.
XX Homo sapiens.
OS US607879-B1.
XX 19-AUG-2003.
XX 09-FEB-1998; 98US-00023655.
XX 09-FEB-1998; 98US-00023655.
XX 09-FEB-1998; 98US-00023655.
XX (INCY-) INCYTE CORP.
XX Cocks BG, Stuart SG, Selthamer JJ;
PI WPI; 2003-895307/82.
XX A composition comprising a plurality of cDNAs, useful for detecting
PT altered expression of genes in an immunological response or for
PT diagnosing and treating an immunopathology, e.g. Crohn's disease, asthma
PT or osteoarthritis.
XX Claim 1, SEQ ID NO 689; 50pp; English.
XX The invention relates to a composition comprising a plurality of cDNAs
CC for detecting the altered expression of genes in an immunological
CC response. The invention also relates to a method of diagnosing or
CC monitoring the treatment of an immunopathological condition in a sample,
CC comprising obtaining nucleic acids from a sample, contacting the nucleic
CC acids of the sample with an array comprising the plurality of cDNAs under
CC conditions to form one or more hybridisation complexes, detecting the
CC hybridisation complexes and comparing the levels of the detected
CC hybridisation complexes with the level of hybridisation complexes
CC detected in a non-diseased sample, where an altered level of the detected
CC hybridisation complexes correlates with the presence of an
CC immunopathological condition. Also disclosed are an expression profile
CC comprising a microarray and a plurality of detectable complexes and a
CC method for identifying a plurality of polynucleotide probes. The cDNAs
CC are useful as hybridisable array elements in a microarray for monitoring
CC the expression of target polynucleotides. The microarray can be used in
CC the diagnosis of an immunopathology, such as Crohn's disease, asthma,
CC ulcerative colitis, hypereosinophilia, irritable bowel syndrome,
CC osteoarthritis, rheumatoid arthritis or acute monocytic leukemia, and in
CC identifying agents for the treatment of the diseases. The microarray may
CC also be used in drug discovery and development, toxicological and
CC carcinogenicity studies, forensics or pharmacogenomics. The composition
CC may also be used in purification of a subpopulation of mRNAs, cDNAs or

CC genomic fragments. This sequence represents a human cDNA of the
CC invention. Note: The sequence data for this patent did not form part of
CC the printed specification but was obtained in electronic format directly
CC from USPTO at seqdata.uspto.gov/sequence.html.
XX
SQ Sequence 1429 BP; 424 A; 311 C; 360 G; 329 T; 0 U; 5 Other;
Query Match 69.4%; Score 1397.8; DB 11; Length 1429;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 1411; Conservative 0; Mismatches 7; Indels 2; Gaps 1;
QY 595 GGGAAAGCCTTTAGTCGTGGCTATCACTTAGTCAACATCAGAAATCATCTGTGAG 654
Db 10 GGGAAAGCCTTTAGTCGTGGCTATCACTTAGTCAACATCAGAAATCATCTGTGAG 69
QY 655 AAACCTTATGAATGTAAAGAAATGTAAAGAGCCTTCGTTGGGGCAATCAGCTTAA 714
Db 70 AAACCTTATGAATGTAAAGAAATGTAAAGAGCCTTCGTTGGGGCAATCAGCTTAA 129
QY 715 CATCAAAAATTCATACCTGGGAGAGAGCCTTACGAATGTAAAGACTGTGGG--AAGGCTT 772
Db 130 CATCAAAAATTCATACCTGGGAGAGAGCCTTACGAATGTAAAGACTGTGGGAGAGCTT 189
QY 773 TTGCATGGGGCTCAAGCCTCGTTATTTCAATAGAGATTCAATCTGTGAAAAAACCCTATG 832
Db 190 TTGCATGGGGCTCAAGCCTCGTTATTTCAATAGAGATTCAATCTGTGAAAAAACCCTATG 249
QY 833 AATGTAAAGACTGTGAAAGAGCCTTTGGCGGTGTATGAGCTCAAGACACAGAGAT 892
Db 250 AATGTAAAGACTGTGAAAGAGCCTTTGGCGGTGTATGAGCTCAAGACACAGAGAT 309
QY 893 TCACACCTGGGGAGAAAGACTTACGAATGTGGAAGAGCCTTTAGCGGTGT 952
Db 310 TCACACCTGGGGAGAAAGACTTACGAATGTGGAAGAGCCTTTAGCGGTGT 369
QY 953 ATTAACCTTATTCAGACAGAGAAATTTCAATGTGGGGAGAGCCTTACAGTGTAAAGACT 1012
Db 370 ATTAACCTTATTCAGACAGAGAAATTTCAATGTGGGGAGAGCCTTACAGTGTAAAGACT 429
QY 1013 GTGGGAAGGCTTTTATTTGTGGTTCAAGCCTTATTCAGCATTAAGAAATTCACAGAGTG 1072
Db 430 GTGGGAAGGCTTTTATTTGTGGTTCAAGCCTTATTCAGCATTAAGAAATTCACAGAGTG 489
QY 1073 AGAAACCTTATGAATGTAAAGAAATGTGGAAGAGCCTTTATCTGATCAATTACTTACTC 1132
Db 490 AGAAACCTTATGAATGTAAAGAAATGTGGAAGAGCCTTTATCTGATCAATTACTTACTC 549
QY 1133 AGCATCAGAAATGTCAACACCGGTGGAAGCCTTACAGATGTAAAGAGTGTGGGAAGGCTT 1192
Db 550 AGCATCAGAAATGTCAACACCGGTGGAAGCCTTACAGATGTAAAGAGTGTGGGAAGGCTT 609
QY 1193 TTGCGTGGGGTTTCGAGCCTCGTTTAAAGACAGAGAGATCATACGGGCGAGAAGCCGTACA 1252
Db 610 TTGCGTGGGGTTTCGAGCCTCGTTTAAAGACAGAGAGATCATACGGGCGAGAAGCCGTACA 669
QY 1253 AGTCACAGAAATGTGGAAGAGCCTTCAATTTGTGCTTATCACTTCAACAGACAGAGAA 1312
Db 670 AGTCACAGAAATGTGGAAGAGCCTTCAATTTGTGCTTATCACTTCAACAGACAGAGAA 729
QY 1313 TCCACAGAGGGGAAACCCCGTAAATGTAAAGAGTGTGGGAAGGCTTCAATTTAGAT 1372
Db 730 TCCACAGAGGGGAAACCCCGTAAATGTAAAGAGTGTGGGAAGGCTTCAATTTAGAT 789
QY 1373 CGAGCTGTGAAACATGAGAAATTCATACCGGGGTGAAAACCTTATGGGTGTACAGAT 1432
Db 790 CGAGCTGTGAAACATGAGAAATTCATACCGGGGTGAAAACCTTATGGGTGTACAGAT 849
QY 1433 GTGGGAAGGCTTTAGTGAACGGGCTTACGCTTACCAATCAGAAAGGCACTGTGGG 1492
Db 850 GTGGGAAGGCTTTAGTGAACGGGCTTACGCTTACCAATCAGAAAGGCACTGTGGG 909
QY 1493 CGAAATCTTACGAATGTAAAGAGTGTGGGAAGGCACTGTAAACCAATCTCCGAG 1552

Db	910	CGAAATCTCAAGATGTGTAAGGAGTGGGGAAGGCATGTAAACCACTTAACCATCTCCGAG	969
Qy	1553	AACATCAGAGGATCCACCAACGTTGAAGAGCCCTTTTGAAGCGAGTAGCCGCTCGTATCT	1612
Db	970	AACATCAGAGGATCCACCAACGTTGAAGAGCCCTTTTGAAGCGAGTAGCCGCTCGTATCT	10529
Qy	1613	ATGGTTTCGCTTCCACAGTTGTGTATCCCTGAGTCAATGAGTGTCAAAAATTTAATG	16722
Db	1030	ATGGTTTCGCTTCCACAGTTGTGTATCCCTGAGTCAATGAGTGTCAAAAATTTAATG	10899
Qy	1673	GAAATTCACAGAAATAAAGAAATTTTAAGTCTCAATGCTGTGCCCCCTTCTGAGTACGTGA	1732
Db	1090	GAAATTCACAGAAATAAAGAAATTTTAAGTCTCAATGCTGTGCCCCCTTCTGAGTACGTGA	11449
Qy	1733	TGAAATCTCTGCGCTGCTCGGCTCCAGCCGCGCGGGAGTGAATCATCCCTTGGTCCAGC	17922
Db	1150	TGAAATCTCTGCGCTGCTCGGCTCCAGCCGCGCGGGAGTGAATCATCCCTTGGTCCAGC	12099
Qy	1793	ACATCCACGCTGTATACGCCACCAACCTCGTATGACTTATAGTACCGCTTGTGTATCA	18522
Db	1210	ACATCCACGCTGTATACGCCACCAACCTCGTATGACTTATAGTACCGCTTGTGTATCA	12659
Qy	1853	GATCACTATCTCCACAGCATCAAGTGCCTGTGTGCCAAGTAGTCTTCACTTGTCTTAACGT	19122
Db	1270	GATCACTATCTCCACAGCATCAAGTGCCTGTGTGCCAAGTAGTCTTCACTTGTCTTAACGT	13299
Qy	1913	GGCCCCAGAGAGCAGAGTAGTGTATGCTGTGATTCGGATATGCAAGAGAAAGCCACAA	19722
Db	1330	GGCCCCAGAGAGCAGAGTAGTGTATGCTGTGATTCGGATATGCAAGAGAAAGCCACAA	13899
Qy	1973	AGTGTCTCTTTTAATGAAAAAGGTGAAGTTCTCACTT	2012
Db	1390	AGTGTCTCTTTTAATGAAAAAGGTGAAGTTCTCACTT	1429

XX	RESULT. 11
XX	AD883430
XX	ID AD883430 standard; cDNA, 1429 BP.
XX	AC
XX	AD883430;
XX	DT 11-AUG-2005 (first entry)
XX	DE Human lymph node cDNA #689.
XX	SS gene; human; immunological response; blood cell; cancer;
KW	immunopathological; AIDS; allergy; anaemia; asthma; atherosclerosis;
KW	bronchitis; ulcerative colitis; diabetes; multiple sclerosis;
KW	osteoporosis; pancreatitis; infection; arthritis; lymph node.
XX	
OS	Homo sapiens.
XX	
PN	US2004077003-A1.
XX	
PD	22-APR-2004.
XX	
PF	14-AUG-2003; 2003US-00641643.
XX	
PR	09-FEB-1998; 98US-00023655.
XX	
PA	(INCY-) INCYTE CORP.
XX	
PI	Cocks BG, Stuart SG, Selhammer JF;
XX	
DR	WPI; 2004-387937/36.
XX	
PT	New compositions having a number of first, second and third
PT	polynucleotide probes, useful in research and diagnostic applications in
PT	cancer and immunological conditions e.g. AIDS, diabetes, osteoporosis and
PT	infections.
XX	
XX	Claim 15; SEQ ID NO 689; 16pp; English.
XX	

CC The invention relates to polynucleotides which are used as probes to
CC detect genes differentially expressed in an immunological response,
CC abundantly expressed in an immunological response and/or coding for a
CC polypeptide known to regulate blood cell biology. The polynucleotides are
CC useful in research and diagnostic applications particularly in cancer and
CC immunopathological conditions, such as AIDS, allergies, anaemia, asthma,
CC atherosclerosis, bronchitis, ulcerative colitis, diabetes, multiple
CC sclerosis, osteoporosis, pancreatitis, infections and arthritis. The
CC present sequence represents a human lymph node cDNA used to detect blood
CC cell and immunological response gene expression. Note: The present
CC sequence does not appear in the printed specification but was obtained in
CC electronic format from the USPTO web site
CC (seqdata.uspto.gov/sequence.html?DocID=20040077003).
XX

Sequence 1429 BP; 424 A; 311 C; 360 G; 329 T; 0 U; 5 Other;
SQ

Query Match	Similarity	59.4%	Score 1397.8	DB 13	Length 1429
Best Local	Similarity	99.4%	Pred. No. of		
Matches 1411	Conservative	0	Mismatches	7	Indels 2
					Gaps 1
QY	595	GGGAAGGCTTTAGTCGTGGCTATCAACTTATGTCACATCAGAAAATTCATCTAGTGAG	654		
Db	10	GGGAAGGCTTTAGTCGTGGCTATCAACTTATGTCACATCAGAAAATTCATCTAGTGAG	69		
QY	655	AAACCTTATGAAATGTAAGAATGTAAGAAGGCTTCCTGGGGCATCAGCTTACCA	714		
Db	70	AAACCTTATGAAATGTAAGAATGTAAGAAGGCTTCCTGGGGCATCAGCTTACCA	129		
QY	715	CATCAAAAAATTCATCTGGGGAGAAAGCCCTCAGATGTAAAGACTGTGG--AAGCTT	772		
Db	130	CATCAAAAAATTCATCTGGGGAGAAAGCCCTCAGATGTAAAGACTGTGGGAAGGCTT	189		
QY	773	TTGATGGGGCTCAAGCCTCGTTATTCATAAGAGATTCATCTGTGTGAAAAACCTATG	832		
Db	190	TTGATGGGGCTCAAGCCTCGTTATTCATAAGAGATTCATCTGTGTGAAAAACCTATG	249		
QY	833	AATGTAAAGATGTGGAAAGGCTTTGGCGGTGTATAGCTACTCAGACAGAGAT	892		
Db	250	AATGTAAAGATGTGGAAAGGCTTTGGCGGTGTATAGCTACTCAGACAGAGAT	309		
QY	893	TCCACTCTGGGGAAGAAAGACTCAGATGCAAGAAGCTGTGGAAAGCTTTAGCCGTGTG	952		
Db	310	TCCACTCTGGGGAAGAAAGGNTAGCAATGCAAAAGACTGTGGAAAGCTTTAGCCGTGTG	369		
QY	953	ATTAACCTTATTCAGACAAAGAAATTCATAGTGGGGAGAAAGCTTACAGAGTATAAGCT	1012		
Db	370	ATTAACCTTATTCAGACAAAGAAATTCATAGTGGGGAGAAAGCTTACAGAGTATAAGCT	429		
QY	1013	GTGGGAAGGCTTTATTTGTGGTCAAGCCCTTCACAGCATTAAGAATTCACACAGTG	1072		
Db	430	GTGGGAAGGCTTTATTTGTGGTCAAGCCCTTCACAGCATTAAGAATTCACACAGTG	489		
QY	1073	AGAAACCTTATGAAATGTCAGAATGTGGAAAGGCTTTACTCGAGTCATTAACCTTACTC	1132		
Db	490	AGAAACCTTATGAAATGTCAGAATGTGGAAAGGCTTTACTCGAGTCATTAACCTTACTC	549		
QY	1133	AGCATCAAAATTCACACCGGTGAAGAAAGCTTCAGAAATGAAGAAGTGGGAAGGCT	1192		
Db	550	AGCATCAAAATTCACACCGGTGAAGAAAGCTTCAGAAATGAAGAAGTGGGAAGGCT	609		
QY	1193	TTGCGTGGGGTTCGAGCCTCGTTAAGCAGAGAGATACATACGGCGAGAAAGCCGTACA	1252		
Db	610	TTGCGTGGGGTTCGAGCCTCGTTAAGCAGAGAGATACATACGGCGAGAAAGCCGTACA	669		
QY	1253	AGTGCACGAATGTGGGAAGGCTTCAATTGTGCTATCACTCAGACGAGAGAA	1312		
Db	670	AGTGCACGAATGTGGGAAGGCTTCAATTGTGCTATCACTCAGACGAGAGAA	729		
QY	1313	TCACACAGGGGAAACCCCGTATAAATGTAGAAGTGTGGGAAGGCTTCAATTATGAT	1372		
Db	730	TCACACAGGGGAAACCCCGTATAAATGTAGAAGTGTGGGAAGGCTTCAATTATGAT	789		
QY	1373	CGAGCGCTGTGAACATGAGAAATTCATACGGGGGTGAACCCGTATGGTGTACAGAT	1432		

```

Db 790 CGAGCTCGTGAATGAGAAATTCATCCGGGTGAAACCTATGGGTTCACAGAT 849
Qy 1433 GTGGGAAGAGCTTTAGTCACGGCCATCAGCTTACCAACATCAGAAAAAGCAAGTGGG 1492
Db 850 GTGGGAAGAGCTTTAGTCACGGCCATCAGCTTACCAACATCAGAAAAAGCAAGTGGG 909
Qy 1493 CGAAATCTCAAGATGTAAGAGTGGGGAAGGCATGTATCACTTAACCATCTCCAG 1552
Db 910 CGAAATCTCAAGATGTAAGAGTGGGGAAGGCATGTATCACTTAACCATCTCCAG 969
Qy 1553 AACATCAGAGATCCACACAGATTGAAGAGCCCTTTGAACGAGTAGCCCGCTCGATCT 1612
Db 970 AACATCAGAGATCCACACAGATTGAAGAGCCCTTTGAACGAGTAGCCCGCTCGATCT 1029
Qy 1613 ATGTTTGGCTTCCACAGATTGTATCCTGAGTCAACTGACAGTTCAAAAATATTAAATG 1672
Db 1030 ATGTTTGGCTTCCACAGATTGTATCCTGAGTCAACTGACAGTTCAAAAATATTAAATG 1089
Qy 1673 GAAATTTCCAGAAATTAAGAAATTTTAAGTCTCAATGGTGTGCTTGTGAGTAGCGTGA 1732
Db 1090 GAAATTTCCAGAAATTAAGAAATTTTAAGTCTCAATGGTGTGCTTGTGAGTAGCGTGA 1149
Qy 1733 TGAATTTCTGCTGCTCGGGCTCCAGCCGGCCGGGGAGTGAAGTCAATCCCTTGTCCAGC 1792
Db 1150 TGAATTTCTGCTGCTCGGGCTCCAGCCGGCCGGGGAGTGAAGTCAATCCCTTGTCCAGC 1209
Qy 1793 ACATCCACGCTGTATACGCCACCAACCTGCTAGTACTTAAAGACCGTCTTGAGTATCA 1852
Db 1210 ACATCCACGCTGTATACGCCACCAACCTGCTAGTACTTAAAGACCGTCTTGAGTATCA 1269
Qy 1853 GATCAACTATCCACGATCAAGTGTCTGTGCCAAGTATGTCTCACTTGTCTTAAAGT 1912
Db 1270 GATCAACTATCCACGATCAAGTGTCTGTGCCAAGTATGTCTCACTTGTCTTAAAGT 1329
Qy 1913 GGGCCCGAGAGCAGAGTATGATGCTGTGATTCGATATGCCAAGAGAGAGCCAGCA 1972
Db 1330 GGGCCCGAGAGCAGAGTATGATGCTGTGATTCGATATGCCAAGAGAGAGCCAGCA 1389
Qy 1973 AGTGTCTCTTTTAAATGAAAAGTGAAGTTCTCAACTT 2012
Db 1390 AGTGTCTCTTTTAAATGAAAAGTGAAGTTCTCAACTT 1429

RESULT 12
ADFI4367
ID ADFI4367 standard; DNA; 4483 BP.
XX
AC ADFI4367;
XX
DT 12-FEB-2004 (first entry)
XX
DE Human endometrial-specific DNA - SEQ ID 332.
XX
KW endometrial-specific; ESNAs; ESP; vaccine; cancer;
KW squamous cell carcinoma; endometrium; gene therapy; transgenic; human;
KW ds.
XX
OS Homo sapiens.
XX
PN WO2003059927-A1.
XX
PD 24-JUL-2003.
XX
PF 20-DEC-2002; 2002WO-US041413.
XX
PR 21-DEC-2001; 2001US-0342751P.
XX
PA (DIAD-) DIADEXUS INC.
XX
PI Sun Y, Liu C;
XX
DR WPI; 2003-598500/56.

```

```

XX New endometrial specific genes and proteins, useful in gene therapy or as
PT vaccines for treating endometrial cancer or non-cancerous endometrial
PT diseases, as well as for diagnosing, monitoring or staging such diseases.
PS Claim 1; SEQ ID NO 332; 1156bp; English.
XX
CC The invention relates to a novel isolated endometrial specific nucleic
CC acid molecule (ESNA) comprising a sequence that encodes an endometrial-
CC specific polypeptide (ESP). The molecule of the invention may be useful
CC for generating a vaccine and for identifying, diagnosing, monitoring,
CC staging, imaging and treating endometrial cancer, such as squamous cell
CC carcinoma, as well as non-cancerous disease states in the endometrium.
CC Furthermore, the polypeptides and polynucleotides may be useful in gene
CC therapy, production of transgenic animals and cells and in the production
CC of engineered endometrial tissue for treatment and research. The current
CC sequence is that of the human endometrial-specific DNA of the invention.
XX
SQ Sequence 4483 BP; 1262 A; 960 C; 1086 G; 1166 T; 0 U; 9 Other;

Query Match 36.5%; Score 734.6; DB 10; Length 4483;
Best Local Similarity 91.2%; Pred. No. 1.9e-221;
Matches 791; Conservative 0; Mismatches 74; Indels 2; Gaps 1;

Qy 47 CAGGCGACATCCTCAGAAAAAGCATCCCGAGAGAGAGACGATCGTTAAACATCTT 106
Db 1184 CAGGCGACATCCTCAGAAAAAGCATCCCGAGAGAGAGACGATCGTTAAACATCTT 1243
Qy 107 AGGTCAAGCTAGCCTCTCGAATTTGTCTTCTTCAAGTGAACCCTCGAAGACTGATC 166
Db 1244 AGGTCAAGCTAGCCTCTCGAATTTGTCTTCTTCAAGTGAACCCTCGAAGACTGATC 1303
Qy 167 AGTCTTCAAGTTCTAATAACATATGAGCCAGGTTTGATGACGTTCCGCGACGTAAGCATAG 226
Db 1304 AGTCTTCAAGTTCTAATAACATATGAGCCAGGTTTGATGACGTTCCGCGACGTAAGCATAG 1363
Qy 227 ACTTTTTCAGAGAGAGAGGCGCTGTGAACTCTGCTCAGAGGACCTGTACTGGGAAG 286
Db 1364 ACTTTTTCAGAGAGAGAGGCGCTGTGAACTCTGCTCAGAGGACCTGTACTGGGAAG 1423
Qy 287 TGATGCTGAGAGACTACAGTATCTTGATCTCACTGATTTGGAGTCAAGATATGAATAA 346
Db 1424 TGATGCTGAGAGACTACAGTATCTTGATCTCACTGATTTGGAGTCAAGATATGAATAA 1483
Qy 347 AGAGTTTACTTACAGAAAAAAACATTCATGAATAAAGGCTTCAAAAGAAATTCAGTA 406
Db 1484 AGAGTTTACTTACAGAAAAAAACATTCATGAATAAAGGCTTCAAAAGAAATTCAGTA 1543
Qy 407 GAAAGAGTAAATCCCTTGGCGGTAACTGATATGTGAAGTACGCTTGAAGACCAACAGC 466
Db 1544 GAAAGAGTAAATCCCTTGGCGGTAACTGATATGTGAAGTACGCTTGAAGACCAACAGC 1603
Qy 467 GCTCCAGAGGAGATGATCAATCAGATGATCATCAATATATGCAAAAGGCTGTACTA 526
Db 1604 GCTCCAGAGGAGATGATCAATCAGATGATCATCAATATATGCAAAAGGCTGTACTA 1663
Qy 527 GAGAAAGGACCCCTCTTGAACACATCAGACATCATTAAGAGAAATCTTTGAATGTA 586
Db 1664 GAGAAAGGACCCCTCTTGAACACATCAGACATCATTAAGAGAAATCTTTGAATGTA 1723
Qy 587 AGGACTGTGGGAAGGCTTTTACTGCTGCTATCACTTAAGTCAACATCAGAAATTCATA 646
Db 1724 AGGACTGTGGGAAGGCTTTTACTGCTGCTATCACTTAAGTCAACATCAGAAATTCATA 1783
Qy 647 CTGTGAGAAACCTTATGATGTAAAGATTAAGAGGCTTCCGTTGGGGCAATCAAC 706
Db 1784 CTGTGAGAAACCTTATGATGTAAAGATTAAGAGGCTTCCGTTGGGGCAATCAAC 1843
Qy 707 TTAATCAACATCAAAAATTTACTGCTGGGAGAGCCCTTCAAGATTAAGTGAAGCTGTGGGA 766
Db 1844 TTAATCAACATCAAAAATTTACTGCTGGG--GAAACCTTATGGGTGTGAAGATGTGGGA 1901
Qy 767 AGGCTTTTGAATGGGGCTCAAGCTCGTATTCATTAAGAGATTCATCTGTGAAAAAC 826

```



```

Db 910 ACTGAGCATCAGAAAATCATCTGTGTAAGAAACCTTATGAATGTAAATATGTGAAG 969
QY 1105 GCCTTACTCGAGTCATTAATCTTAAGCATCAGAGATCCACACCGGTGAGAAAGCT 1164
Db 970 GCTTTTGTGGGGCTATCACTTACTGACATCAGATATTTTCAATCTGTGAGAAACCC 1029
QY 1165 CACGAATGAAGAGCTGTGGGAAGGCTTTCGCTGGGTTGACGCTCTGTAAAGACAG 1224
Db 1030 TATGATGAGAGAAATGTGGGAAGGCTTTTAAATGGGATCAAGCTTTATTAACATGAA 1089
QY 1225 AGATATACATACCGGGGAGAGCCGTACAAAGTCACAGATGTGGGAAGGCTTCAATTGT 1284
Db 1090 AGAATTCATCTGTGTGAGAAACCTTATGATGATTAAGATGTGGAAAGGCTTTAGTGT 1149
QY 1285 GGCATACCTGCTCAGACAGAGAAATCCACAGAGGAAACCCGTATTAATGTAAAG 1344
Db 1150 GGCTATACCTTTCTCAACATCAGAAAATCCATCTGTGAGAAACCTTTGAATGTAAAG 1209
QY 1345 GAGTGTGGGAAGGCTTTCTATTATGATGAGCTTGTGAAACATGAGAAATTCATACC 1404
Db 1210 GATGTGGGAAGGCTTTAGTTGGGGTTCAACCTTGTATTAACATGAGAGGTTCAATCT 1269
QY 1405 GGGGTGAAACCTTATGGGTGTACAGATGTGGGAAGGCTTTAGTCACGGCCATCAGCTT 1464
Db 1270 GGTGGAATATCCCATGATGTAAAGATGTGGAAAGACCTTTGTATGTGGTATCACTT 1329
QY 1465 ACACAACATCAGAAAAACGACAGTGGGGGCAATCTTAAGATGTAAGAGTGTGGGAAG 1524
Db 1330 ACTGACATCAGGTATTTTCACTGTGTGAGAAACCTTAAGATGTAAGGAATGTGGGAAG 1389
QY 1525 GCATGTACCACTTAACCATCTCCGAGAACATCAGAGATCCACAACATGTAAGAGCC 1584
Db 1390 GCTTTTAAATTTGTGATCAAGCTTTGTCAACATGAAGAAATCCATACAGGGGAGAAACCC 1449
QY 1585 TTT 1587
Db 1450 TAT 1452

```

RESULT 14
ACA62107
ID ACA62107 standard; cDNA; 2143 BP.

AC ACA62107;

DT 20-AUG-2003 (first entry)

DE cDNA encoding human cartilage zinc finger-2 (CZF-2).

XX Human; cartilage zinc finger-2; CZF-1; CZF-2; chondrogenesis; fracture;
KW tumour; chondrosarcoma; bone formation; chondrogenesis detection;
XX chondrogenesis staging; gene; ss.

OS Homo sapiens.

XX Key Location/Qualifiers
FH 25.1581
FT CDS
FT /tag= a
FT /product= "CZF-2"
FT /note= "Cartilage zinc finger-2"

XX US2003039966-A1.

XX 27-FEB-2003.

XX 14-JUN-2001; 2001US-00881578.

XX 14-JUN-2001; 2001US-00881578.

XX (HERI/) HERING T M.
XX (JOHN/) JOHNSTONE B.

PI Hering TM, Johnstone B;
DR WPI: 2003-492160/46.
DR P-PSDB; AB010395.
XX
XX
PT Novel isolated CZF-1 or CZF-2 protein or its variants, useful as an
PT immunogen for preparing an antibody which is immunoreactive with the
PT protein.
XX
XX
XX Claim 12; Fig 5A-C; 35dp; English.

CC The invention describes an isolated CZF-1 or CZF-2 protein (I) or its
CC variants. The polynucleotide (II) encoding (I) or an anti-(I)-antibody
CC are useful for determining the extent of chondrogenesis in a cell or for
CC ascertaining the presence of cells having characteristics of
CC chondrocytes, in a tissue sample obtained from a tissue or fracture
CC cells, by contacting the cell or tissue sample, or RNA isolated from the
CC cell or tissue sample with a nucleic acid probe. (II) is useful for
CC detecting and staging chondrogenesis in cells, for detecting and
CC characterising cells obtained from a tumour that is suspected of being a
CC chondrosarcoma or cells that have characteristics of chondrocytes that
CC may be present in a tumour tissue, to amplify or probe RNA obtained from
CC mesenchymal cells that are being used as screening tools for compounds
CC that are designed to accelerate or retard chondrogenesis, to probe cells
CC obtained from any area of the body where chondrogenesis is suspected of
CC taking place e.g. an area in the body where new bone formation is taking
CC place after a fracture, and to identify chromosomal aberrations to
CC compare with endogenous DNA sequences in patients to identify potential
CC genetic disorders. An anti-(I)-antibody is useful as for identifying
CC tissues that contain normal or elevated levels of (I) and for purifying
CC (I) by affinity chromatography, for identifying and diagnosing diseases
CC associated with elevated or reduced levels of (I), for monitoring the
CC effect of therapeutic agents on the synthesis of (I) by cells in vitro
CC and in vivo and for identifying other proteins, activators and inhibitors
CC which bind to (I). This sequence encodes human cartilage zinc finger-2
CC (CZF-2), a marker for detecting and staging chondrogenesis

XX Sequence 2143 BP; 735 A; 331 C; 478 G; 599 T; 0 U; 0 Other;

Query Match 34.9%; Score 703; DB 8; Length 2143;
Best Local Similarity 69.7%; Pred. 1.3e-211;
Matches 1006; Conservative 0; Mismatches 410; Indels 27; Gaps 3;

```

QY 172 TTCAGTTCTMAAACATATGCCCCAGGCTTTGGTACCTTCCGACGTACCTAATACCTT 231
Db 10 TCGAATTCAGAACCATCATCATGATGGTGTGATTCAGGATGTGGCCATTCACCTTC 69
QY 232 TCTCAGAGAGAGTGGGCGTCTGAACTCTGCTCAGAGGACCTGTATCTGGGACGTGATG 291
Db 70 TCTCAGAGAGAGTGGGAAATGCTGACCTCTGCTCAGAGGACCTTGTACGTGATGATG 129
QY 292 CTGAGAACATCAGTACTGTCTCACTGTATTTGAGTGC---AGCATATGAAATTAAG 348
Db 130 TTGAGGAATATATGTAATCTGTGTGTCACTGATTTGGATGTAAACGATAGAGCAAAA 189
QY 349 AGTTTACTACAGAAAAAAACATTCATGAATTAAGGCTTCCAAAAGGAATTCAGATAGA 408
Db 190 AAAATATTTTTCAGAAAATGATATTTTGAATAAATTTTCCAGCTGGGAGATGAAGGAC 249
QY 409 AGAAGTAATCCCTTGGCC-----GTAATGATATATGAAGGT 447
Db 250 AAAAGTAATAACCTTGGCTTGGAGCATTCATTCAGAAATATATGGAAGGCAAAAAGC 309
QY 448 ACGCTTGAAGACACACAGCGCTCCAGAGGAGTATGTCAATCAGATGATCATCAATATAT 507
Db 310 ATATTGAGGGAGCTTAAAGGACATCAAGAGGATATCTTCAATGATATATGAGTAT 369
QY 508 GTCAAAAGGCTGTCTACTAGAGAGGACCCCTCTTGAACATCAAGAGA---CATCAT 564
Db 370 GAAAAAATACCTTTTCAAGAAAAAGTAATCTTCACTCCACATCAAAAGAAATTCATAT 429
QY 565 AAGAGAAATTCCTTGAATGTAAGGACTGTGGGAAGGCTTTAGTGTGCTATCAACTT 624

```

```

Db      430 ACAGAGAAATCCTATGTTTGAAGAATGTGGAGGCTTGACGTCAGTCCAAACTT 489
Oy      625 AGTCAACATCAGAGAAATCCATACTGGTGAAGAACTTTATGATGTAAGATGTAAGAG 684
Db      490 GTTCAACATGAGAGAACTCATACGCTGAAGAACCTTTGATGTAAGAAATGTGGAAAG 549
Oy      685 GCCTTCCTGTTGGGGCAATCAGCTTACTCAACATCAAAAAATTCATACCTGGGAGAAAGCC 744
Db      550 AATTATTTTAAGTCCATATCACTCAATGTGCATCAGAGATTTTCATCTGTGTGAGAAAGCC 609
Oy      745 TACGAATGTAAGAACTGTGGAGAAAGCTTTTCATGAGGGGCTCAAGCCTCGTTATTCATAG 804
Db      610 TATGAGTGAAGAAATGTGGAGAAAGCTTTTAAAGCTTGTGAGTCAAGCCTTTTAAACATGAG 669
Oy      805 AGGATTCATATCTGTGTAAGAAAGCCCTATGATGTAAGAACTGTGAGAAAGCCCTTCGGCGT 864
Db      670 AGAATTCACATCTGTGTAGAGAAAGCCCTATGATGTAAGAAATGTGGAGAAAGCCCTTTAGT 729
Oy      865 GGTGATGAGCTCACTCAGACCCAGAGATTTCCACATGGGGAGAAAGACTACGAAATGCAAA 924
Db      730 GGCCTATCACCCTTACCAATCAGAAATTCATATTTGTGTGAAATCTTATTAATGTAAAG 789
Oy      925 GACTGTGGAGAAAGCTTTTAAAGCCTGTGTATTAATCTTATTCAGACAAAGAAATTCATAGT 984
Db      790 GAATGTGGAGAAAGCTTTTGTGGGCTCAAGCCTTGTCTAAACATGAGAAATTCATACGA 849
Oy      985 GGGAGAAAGCCTTACAGATGTAAGAACTGTGGAGAAAGCTTTTATTTGTGTGCTCAAGCCTC 1044
Db      850 GGTGAGAAAGCCTTATTAATGTAAGAAATGTGGAGAAAGCCTTCAGTGTGTGCTATCAACTT 909
Oy      1045 ATTGAGCATTAAGAAATTCACACAGGTGAGAAAGCCCTATGATGTCAGAAATGTGGAGAAAG 1104
Db      910 ACTGAGCATCAGAAATTCATATCTGTGTAAGAAAGCTTTATGATGTAAGAAATATGTGAAAG 969
Oy      1105 GCCTTATCTGAGTCAATTAATCTTATCTCAGCATCAAGAAATCCACACCGGTGAGAAAGCCT 1164
Db      970 GCTTTTGTGTGGGCTATCACTTACTCTCAGCATCAAGATTTTATATCTGTGTGAGAAAGCC 1029
Oy      1165 CAGAAATGTAAGAGATGTGGAGAAAGCCTTTCGCTGGGGTTCGAGCCTCGTTAAGACGAG 1224
Db      1030 TATGATGCAAGAAATGTGGAGAAAGCCTTTTAAATTTGGGAGTCAAGTCTTATCAACTGAA 1089
Oy      1225 AGGATACATACGGGGCAGAAAGCCGTACAGTGCAAGATGTGGAGAAAGCCTTCAATTGT 1284
Db      1090 AGAATTCATATCTGTGAGAAAGCCTTATGATGTAAGAAATGTGAGAAAGCCTTTAGTCGT 1149
Oy      1285 GGTATCACCTCTCAGACCGAGAAATTCACACAGGCGCAACCCCGTTAATATGTAAG 1344
Db      1150 GGTATCACCTCTCTCAACATCAAGAAATTCATATCTGTGTAAGAAAGCCTTTGATGTAAAG 1209
Oy      1345 GAGTGTGGAGAAAGCCTTTCATTTATGATCGAGCCTCGTGAAGCAATGAGAAATTCATACC 1404
Db      1210 GAATGTGGAGAAAGCCTTTAGTTGGGTTCAAGCCTTGTAAACATGAGAGATTCATACCT 1269
Oy      1405 GGGGTGAAAGCCTATGGGTGTACAGAAATGTGGAGAAAGCCTTTATGTCACGGCCATCAGCTT 1464
Db      1270 GGTGAGAAATCCATGATGTAAGAAATGTGGAGAAAGCCTTTGTAGTGGGTATCAACTT 1329
Oy      1465 ACACAACTTCAGAAAGCAAGTGGGGCCAAATTCCTACAAATGTAAGAGTGTGGGAAAG 1524
Db      1330 ACTCGATCATCAGGTATTTTCACTGTGTGAGAAAGCCCTATGAATGTAAGAAATGTGGGAAAG 1389
Oy      1525 GCATGTAAACACCTAAACCATCTCCGAGACATCAGAGATTCACAAACAGTTGAAGAGCC 1584
Db      1390 GCTTTTATTTGTGATCAAGCCTTTGTTCAACATGAAGAAATCCATACAGGGAGAAAGCC 1449
Oy      1585 TTT 1587
Db      1450 TAT 1452

```

RESULT 15
ADC30401

```

ID      AD30401 standard; cDNA; 2832 BP.
XX
XX
AC      AD30401;
XX
XX      18-DEC-2003 (first entry)
DT
XX
XX      Human novel cDNA sequence, SEQ ID NO:483.
DE
XX
XX      Human; diagnostic; drug screening; forensics; gene mapping;
KW      biodiversity assessment; Parkinson's disease; Alzheimer's disease;
KW      neurodegenerative diseases; anaemia; platelet disorder; wound; burns;
KW      ulcers; osteoporosis; autoimmune disease; cancer;
KW      molecular weight marker; food supplement; antiparkinsonian; nootropic;
KW      neuroprotective; antianemic; anticoagulant; thrombolytic; vulnerary;
KW      anticancer; osteopathic; immunosuppressive; antiinflammatory; cyostatic;
KW      gene therapy; chromosome 19; gene; ss.
XX
OS      Homo sapiens.
XX
XX      WO2003029271-A2.
PN
XX
XX      10-APR-2003.
PD
XX
XX      24-SEP-2002; 2002WO-US030474.
PF
XX
XX      24-SEP-2001; 2001US-0324631P.
PR
XX
XX      (HYSE-) HYSEQ INC.
PA
XX
XX      Tang TY, Zhang J, Ren F, Xue AJ, Zhao QA, Wang J, Wehrman T;
PI      Zhou P, Ghosh M, Wang D, Ma Y, Asundi V, Wang Z, Weng G;
PI      Haley-Vicente D, Drmanac RT;
XX
XX      WPI; 2003-371981/35.
DR
XX
XX      P-PSPB; AD3031372.
PT
XX
XX      New polynucleotide and polypeptide useful for diagnosing, preventing or
PT      treating conditions such as neurodegenerative diseases, anemias, platelet
PT      disorders, wounds, burns, ulcers, osteoporosis, autoimmune diseases or
PT      cancer.
XX
XX      Claim 1, SEQ ID NO 483; 1185bp; English.
PS
XX
XX      The invention relates to 971 novel human cDNA sequences (ADC39919-
XX      ADC30889) and the polypeptides they encode (ADC30890-ADC31860). The
XX      invention also relates to nucleic acid sequences over 99% identical with
XX      the novel human cDNAs. The invention additionally encompasses expression
XX      vectors and host cells comprising a nucleic acid of the invention; the
XX      recombinant production of a polypeptide of the invention; an antibody
XX      against a polypeptide of the invention; a method of detecting
XX      polynucleotides or polypeptides of the invention; and methods of
XX      identifying a compound which binds to a polypeptide of the invention. The
XX      invention further discloses methods of preventing, treating or
XX      ameliorating a medical condition; kits comprising polynucleotide probes
XX      and/or monoclonal antibodies for carrying out the methods of the
XX      invention; methods for the identification of compounds that modulate the
XX      expression or activity of the polynucleotide and/or polypeptide; and 767
XX      contig sequences corresponding to the cDNA sequences of the invention
XX      (ADC31861-ADC32627) and the polypeptides encoded by the contigs (ADC32628
XX      -ADC33394). The nucleic acids and polypeptides of the invention are
XX      useful in diagnostics, drug screening, forensics, gene mapping, in the
XX      identification of mutations responsible for genetic disorders or other
XX      traits, for assessing biodiversity, and in producing many other types of
XX      data and products dependent on DNA and amino acid sequences. They are
XX      also used for treating diseases such as Parkinson's disease, Alzheimer's
XX      disease and other neurodegenerative diseases, anaemia, platelet
XX      disorders, wounds, burns, ulcers, osteoporosis, autoimmune diseases or
XX      cancer. The nucleic acids may also be used as hybridisation probes or
XX      primers, and in the recombinant production of a protein. The polypeptides
XX      are also useful in generating antibodies, as molecular weight markers,
XX      and as food supplements. The present sequence represents a specifically
XX      claimed human cDNA sequence of the invention. Note: The sequence data for
XX      this patent did not form part of the printed specification, but was

```

CC obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 2832 BP; 915 A; 552 C; 632 G; 733 T; 0 U; 0 Other;

Query Match 30.8%; Score 620.6; DB 10; Length 2832;
Best Local Similarity 65.6%; Pred. No. 2.4e-185;
Matches 964; Conservative 0; Mismatches 479; Indels 26; Gaps 3;

```

QY 145 GGAAACCCGAGAGAGCTGATCAGTTCTT-CAGTTCTAAAACAATGCGCCAGGGTTTGG 202
DB 160 GGAAACCCGAGAGAGAGCTGATCAGTTCTTTCAGACTTAAACCAATGCTCGGAATATAG 219
QY 203 TACAGCTTCCGACGACGATGAGCACTTTCTTCAGAGAGAGTGGGCTCTGAACTCTG 262
DB 220 TGATGTTCCAGGAGTGTGGCATTGACTTCTTCAGAGAGAGTGGGAAATCCCTGGA 279
QY 263 CTCAGAGGAGCCTGTACTGGGACGTGATGCTGGAGAACTACAGTAACTTGGTCTCACTGG 322
DB 280 CTCAGAGAGATTTGTTAGAGATGTGATTTGGAGAACTTAGCAACTTGGTATCACTAG 339
QY 333 ATTTGAGTCAACATATGAAAAAAGAGTTTACCTACAGAAAAAAACATTCATGAAATTA 382
DB 340 ACTTGCTTCAAGGTGTGCAAGTAAAGAACTTATCTCAGAAAAAGAACTTATGAAACAG 399
QY 383 GGGCTTCCAAAAGAAATTCAGATAGAGAAATTAATCCCTTG----- 424
DB 400 AATTAATCCCAATGGGAAATGAGTGAACGACTTGAATAACGTGATCTTGAAGATCCAAAT 459
QY 425 ---GCCGTAACTGATATGTGAAGTACGCTTGAAGAACAACAGGCGCTCAGAGGGAGGT 481
DB 460 CCAGGAGATTTATTTGGAAGCCAAAGGCAAGATGAGAGACACAGAAAAATTCAGAAAGAA 519
QY 482 ATGTCAATCAAGTATCATCAATTAATGTCAAAAGGCTGCTACTAGAGAGGACCCCTC 541
DB 520 ATTTGAGGCAAGGATGATCATATATGACAAATGTCCATTTTCAACCAAGCAATCTACT 579
QY 542 CTAGAAACATCAAGAGA---CATCATAGAGAGAAATTCCTTGAATGTAAGACTGTGGGA 598
DB 580 TATCTCAACATTCAGAGATGTCAATCTGAGAAACCTTATTAAGTAAAGTAAAGTGGGA 639
QY 599 AGGCTTTAGTGTGCTATCAACTTATGCAATCAGAAAAATCCATCTGGTGAAGAAC 658
DB 640 AAGCTTTCAGAGAGAGCTCAACCTAACCAATCAATTAATGTTACTGCTGTAAGAAC 699
QY 659 CTTATGAATGTAAGAGATGTAAGAGGCTTCGTTGGGCAATCAGCTTACTCAACATC 718
DB 700 CCTATGAATGTAGCAATGCGGGAAGGCTTTAGTGTGATTCAGAACTCAGCTTCAATC 759
QY 719 AAAAAATTCATCTGGGGAAGAGCCCTACGAATGTAAAGCTGTGGAGAGGCTTTTCAT 778
DB 760 AGAGACTTCATCTGTGTGAGAAACCTTATGCAATGTAAAGAAATGTGGAGAGGCTTTTACTC 819
QY 779 GGGGCTCAAGGCTCGTTATTCATTAAGAGGATTCATCTGGTGAATAAAACCTTATGATGA 838
DB 820 AAGCTCACAATTTATTTACATCATAGAAATTCATCTGTGAAAAACCATTAATATGTG 879
QY 839 AAGACTGTGAAAGGCTTTCGCGCTGTGATGAGCTCACTCAGACCCAGAGATTCACA 898
DB 880 AAGAAATGTGGAAGGCTTTATTCGTAGCTCAAACTTACCAGATCAAAAAAGTTCAATA 939
QY 899 CTGGGGAAGAAACATACGAATGCAAGAACTGTGGGAAGACTTTAGCCGTGTATTAAC 958
DB 940 CTGGGTGAGAAACCTTAATGATGTAAAGAAATGTGGGAAGGCTTTTACTCAGAAATTCACAAC 999
QY 959 TTATTACAGCAAGAATTCATAGTGGGAGAGAGCTTACAGATGTAAAGACTGTGGGA 1018
DB 1000 TTACACTACACAGAGACTTCACTGCTGTAAGAAAGCTCATGATTAAGATTAAGATGAGGA 1059
QY 1019 AGGCTTTTATTTGTGTTCAAGCTTCATTCAGCATTAAGAAATTCACACAGGTGAAGAAC 1078
DB 1060 AGGTCTTACTACGCTCTCACAACATTAATCTGATTAAGAAATTCATACCGGTGAGAAAC 1119

```

```

QY 1079 CCTATGAATGTCAAGATGTGGGAAGGCTTACTCGAGTCAATTAACCTTACTCAGACATC 1138
DB 1120 CCTATGAATGTAGGAATGTGAAAAAGCTTTTATTTGTGCTCAGACGCTTCTCAACATC 1179
QY 1139 AGAAGATCCACACCCGTAAGAGAGCTTCAAGATGTAAAGAGTGTGGGAAGGCTTTCGCT 1198
DB 1180 AGAAATTCATTAATGTGGGAAAAACCATATGATGTAAAGAAATGTGGAAGGCTTTATATC 1239
QY 1199 GGGGTTCAAGCTTCTTAAACAGAGAGATACATACGGGCGAGAAAGCCGTACAAATGCA 1258
DB 1240 GGGGCTCACTACTGTGGAACATCAGAGATTCATATCTGTGAAAAACCTTAAATATGTG 1299
QY 1259 CAGAAATGTGGGAAGGCTTCAATGTGCTATCACTCACTCAGACGAGAGATTCACA 1318
DB 1300 AAGAAATGTGGGAAGGCTTATCCGTGCTCAAACTTATCAACACAGAAATTCACA 1359
QY 1319 CAGCGAAACCCGTAATAATGTAAAGAGTGTGGAGGCTTTCATTTATGATCGAGCC 1378
DB 1360 CCAATGAAAAAGCCATATGATGTAAAGAAATGTGAAAGATGTTTATGATATGCTCAACAC 1419
QY 1379 TCGTAAACATGAGAAATTCATACCGGGGTGAACCTTATGGGTGTACAGATGTGGGA 1438
DB 1420 TTAATCAACATCAGAGAAATACACACTGTGAGAAACCTTATCAATGTAAAGAAATGTGGA 1479
QY 1439 AGAGCTTATGTCAGGCGCATCAGTTACACAACATCAGAAAAAGCAAGTGGGGGAAT 1498
DB 1480 AAGCTTTAATGTGTGCTCACTCTTACACGACACAGAGGTTTCAATCTGTGTGAGAAC 1539
QY 1499 CCTAGAAATGTAAAGAGTGTGGGAAGGATGTAAACCTTAACATCTCCGAGAACATC 1558
DB 1540 CCTATGAATGTAAAGAAATGTGAAAAACCTTATAGTGTGTGCTCAGAACTTACTCAACATG 1599
QY 1559 AGAGGATCCACAACAGTTGAAGAGCTTT 1587
DB 1600 AGCGAATTCACACAGGTGAGAAACCTTAT 1628

```

Search completed: February 6, 2006, 16:26:11
Job time : 1231 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Biocelleraion Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 6, 2006, 16:05:51 : Search time 10230 Seconds
(without alignments)
11196.437 Million cell updates/sec

Title: US-10-111-257-3

Perfect score: 2015

Sequence: 1 cgccttcgaccggtgacgc.....gtgaagttcacttaaa 2015

Scoring table: IDENTITY_NUC

Searched: 5883141 seqs, 28421725653 residues

Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

GenEmbl:*

1: gb_ba:*
2: gb_in:*
3: gb_env:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pr:*
9: gb_ro:*
10: gb_scs:*
11: gb_sy:*
12: gb_un:*
13: gb_vl:*
14: gb_htg:*
15: gb_pl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2015	100.0	2015	6	AX113689 Sequence
2	2015	100.0	2015	8	AF272148 Homo sapi
3	1986.8	98.6	2098	6	AX427523 Sequence
4	1967.4	97.6	2173	6	AX274843 Sequence
5	1958	97.2	2169	6	BD127591 Primer fo
6	1958	97.2	2169	6	CO783329 Sequence
7	1958	97.2	2169	8	AK075226 Homo sapi
8	1768	87.7	2010	6	AX405913 Sequence
9	1690.8	83.9	138579	8	AC011487 Homo sapi
10	1659.4	82.4	1904	8	AF251515 Homo sapi
11	1645.4	81.7	2838	8	AK128736 Homo sapi
12	1619.4	80.4	2196	6	CS118206 Sequence
13	1619.4	80.4	2196	8	BC009433 Homo sapi
14	1469.6	72.9	219392	8	AC146884 Callitric
15	1408.8	69.9	1420	6	CQ718116 Sequence
16	1397.8	69.4	1420	6	AR380144 Sequence
17	1091.8	54.2	2148	4	AY375188 Canis fam
18	1063.4	53.8	1084	8	BC009365 Homo sapi

19	802.2	39.8	236051	14	AC160802 Bos tauru
20	703.8	34.9	1910	6	CO728249 Sequence
21	703	34.9	2143	6	AB360880 Sequence
22	698.8	34.7	2789	8	AB066540 Macaca fa
23	620.6	30.8	2990	8	BC026210 Homo sapi
24	618.6	30.7	95480	8	AC011508 Homo sapi
25	618	30.7	1945	6	AX714436 Sequence
26	618	30.7	1945	6	AK056753 Homo sapi
27	617.2	30.6	265528	14	AC159553 Bos tauru
28	615.4	30.5	2235	8	AX835235 Sequence
29	615.4	30.5	2235	8	AK098175 Homo sapi
30	613.8	30.5	1664	8	AY500359 Homo sapi
31	611	30.3	1965	6	CO731631 Sequence
32	607.2	29.7	218270	14	AC079499 Mus muscu
33	599.4	29.7	1245	8	AY500364 Homo sapi
34	573.8	28.5	1709	8	BC101333 Homo sapi
35	564.4	28.0	2248	6	CO725766 Sequence
36	558	27.7	2314	6	CO730857 Sequence
37	546	27.1	1335	6	CO739626 Sequence
38	546	27.1	3163	8	BC032863 Homo sapi
39	546	27.1	173669	14	AC017110 Homo sapi
40	546	27.1	179203	8	AC009194 Homo sapi
41	546	27.1	189486	8	AC008733 Homo sapi
42	546	27.1	201981	14	AC073640 Homo sapi
43	544.4	27.0	1970	8	HSHZP10
44	544.4	27.0	2485	8	AK094146 Homo sapi
45	538.6	26.7	1813	6	AX274897 Sequence

ALIGNMENTS

RESULT 1
LOCUS AX113689 2015 bp DNA linear PAT 01-MAY-2001
DEFINITION Sequence 3 from Patent WO0127265.
ACCESSION AX113689
VERSION AX113689.1 GI:13939857
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homniidae; Homo.
REFERENCE
1 Bullerdiel, J., Rippe, V., Weiboom, M. and Beilge, G.
Nucleic acid sequences of hyperplasia and tumours of the thyroid
JOURNAL Patent: WO 0127265-A 3 19-APR-2001;
Universitaet Bremen (DE)

FEATURES

source

1. 2015
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
187..1578
/note="unamed protein product; RITA"

CDS

/codon_start=1
/protein_id="CAC37860.1"
/db_xref="GI:13939858"
/translation="MAOGLVFPADVAIDFSEWACINSAORDLYMDVLMENYSILVS
LDLSAYBNKSLPTKKNHETBARSKSDRSKSLGNWTCGLTEPRLRGRGVNO
MITNYVPRATRGTPPTTRHORHKHNSFEKDCGKAFRSRYQSLQSHOKITGKPYE
CKECKAFRWENQLOHOKHTGKPYECKDCGKAFRWGSLVHKRIHTEKPYECK
DCGKAFRWGDELTOHOKHTGKPYECKDCGKAFRWGSLVHKRIHTEKPYECK
GKAFICGSLIOHOKHTGKPYECKDCGKAFRWGSLVHKRIHTEKPYECKGK
AFRWGSLVHKRIHTEKPYECKDCGKAFRWGSLVHKRIHTEKPYECKGKAF
IYGSILVHERIHTGVKPYECKGKAFRWGSLVHKRIHTEKPYECKGKAF
LNIHREHORHINS"

ORIGIN

Query Match 100.0%; Score 2015; DB 6; Length 2015;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2015; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCTTCTGCACCGGTGACCGCAACCGCTGTGTCTCCGCCAGTCCGCGCAGGCCAGCATCTT 60
Db 1 CGCTTCTGCACCGGTGACCGCAACCGCTGTGTCTCCGCCAGTCCGCGCAGGCCAGCATCTT 60
QY 61 TCAGAAAAAGCATCCCCGAGAGAGAGAGAGAGATCGTTAAACATCTTAAAGTCAAGCTCTAC 120
Db 61 TCAGAAAAAGCATCCCCGAGAGAGAGAGAGAGATCGTTAAACATCTTAAAGTCAAGCTCTAC 120
QY 121 CTCTGGAAATTTGTCTTCTTCAAGTGAACCCCGAGAGAGATGATGATGATGATGATGATGAT 180
Db 121 CTCTGGAAATTTGTCTTCTTCAAGTGAACCCCGAGAGAGATGATGATGATGATGATGATGAT 180
QY 181 AAAACATAGTGGCCAGGGGTTTGTGAGCTTCCGCGAGTACGATGATGATGATGATGATGATGAT 240
Db 181 AAAACATAGTGGCCAGGGGTTTGTGAGCTTCCGCGAGTACGATGATGATGATGATGATGATGAT 240
QY 241 GAGTGGGCTGTCTGTAACTCTGCTCAGAGGAGCTGTAACTGGAGCTGATGCTGAGAGAC 300
Db 241 GAGTGGGCTGTCTGTAACTCTGCTCAGAGGAGCTGTAACTGGAGCTGATGCTGAGAGAC 300
QY 301 TACAGTAACTTGTGTCTCACTGATTTGAGTCAATGATGATGATGATGATGATGATGATGATGAT 360
Db 301 TACAGTAACTTGTGTCTCACTGATTTGAGTCAATGATGATGATGATGATGATGATGATGATGAT 360
QY 361 GAAAAAATCATTCATGATAAATGAGGCTTCCAAAAGAGATTCAGATGAGAGATGATGATGATGAT 420
Db 361 GAAAAAATCATTCATGATAAATGAGGCTTCCAAAAGAGATTCAGATGAGAGATGATGATGATGAT 420
QY 421 CTTCGCGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 480
Db 421 CTTCGCGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 480
QY 481 TATGTCAATCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 540
Db 481 TATGTCAATCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 540
QY 541 CCTAGAACATCAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 600
Db 541 CCTAGAACATCAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 600
QY 601 GCTTTAGT 660
Db 601 GCTTTAGT 660
QY 661 TATGAAATGTAAGATGTAAGATGTAAGATGTAAGATGTAAGATGTAAGATGTAAGATGTAAGAT 720
Db 661 TATGAAATGTAAGATGTAAGATGTAAGATGTAAGATGTAAGATGTAAGATGTAAGATGTAAGAT 720
QY 721 AAAAATCATCTGAGGAGAGAGCCCTAGATGTAAGATGTAAGATGTAAGATGTAAGATGTAAGAT 780
Db 721 AAAAATCATCTGAGGAGAGAGCCCTAGATGTAAGATGTAAGATGTAAGATGTAAGATGTAAGAT 780
QY 781 GGGTCAAGCTGCTTATTCATTAAGAGATTCATCTGTGTGAAAAACCTTAATGATGTAAGAT 840
Db 781 GGGTCAAGCTGCTTATTCATTAAGAGATTCATCTGTGTGAAAAACCTTAATGATGTAAGAT 840
QY 841 GACTGTGAAAGGCTTTGGGCGGTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 900
Db 841 GACTGTGAAAGGCTTTGGGCGGTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 900
QY 901 GGGGAGAGAGACTACGATGCAAAAGACTGTGGAGAGACTTTAGCCGTGTGTAAACTT 960
Db 901 GGGGAGAGAGACTACGATGCAAAAGACTGTGGAGAGACTTTAGCCGTGTGTAAACTT 960
QY 961 ATTCAACAAGAGATTCATTAAGTGGGAGAGAGCTTAAAGATGTAAGATGTAAGATGTAAGAT 1020
Db 961 ATTCAACAAGAGATTCATTAAGTGGGAGAGAGCTTAAAGATGTAAGATGTAAGATGTAAGAT 1020
QY 1021 GCTTTATTTGT 1080
Db 1021 GCTTTATTTGT 1080

QY 1081 TATGAAATGTAAGATGTAAGATGTAAGATGTAAGATGTAAGATGTAAGATGTAAGATGTAAGAT 1140
Db 1081 TATGAAATGTAAGATGTAAGATGTAAGATGTAAGATGTAAGATGTAAGATGTAAGATGTAAGAT 1140
QY 1141 AAGATTCACACCGGTGAGAGAGCTTCAAGATGTAAGATGTAAGATGTAAGATGTAAGATGTAAGAT 1200
Db 1141 AAGATTCACACCGGTGAGAGAGCTTCAAGATGTAAGATGTAAGATGTAAGATGTAAGATGTAAGAT 1200
QY 1201 GGTTCGAGCTTGTATTAAGACAGAGAGATTAACATACGGGCGAGAGAGCTTCAAGATGTAAGAT 1260
Db 1201 GGTTCGAGCTTGTATTAAGACAGAGAGATTAACATACGGGCGAGAGAGCTTCAAGATGTAAGAT 1260
QY 1261 GAATGTGGAGAGGCTTCAATTTGT 1320
Db 1261 GAATGTGGAGAGGCTTCAATTTGT 1320
QY 1321 GGGGAAACCCCGTATTAATGTAAGAGATGTAAGAGATGTAAGAGATGTAAGAGATGTAAGAGAT 1380
Db 1321 GGGGAAACCCCGTATTAATGTAAGAGATGTAAGAGATGTAAGAGATGTAAGAGATGTAAGAGAT 1380
QY 1381 GTGAAACATGAGAGATTCATACCGGGGTGAAACCTTATGGGTGTACAGAAATGTGGAGAG 1440
Db 1381 GTGAAACATGAGAGATTCATACCGGGGTGAAACCTTATGGGTGTACAGAAATGTGGAGAG 1440
QY 1441 AGCTTTAGTCAAGGCTTCACTTACATCAACATCAACATCAACATCAACATCAACATCAACATCA 1500
Db 1441 AGCTTTAGTCAAGGCTTCACTTACATCAACATCAACATCAACATCAACATCAACATCAACATCA 1500
QY 1501 TACGAAATGTAAGAGATGTAAGAGATGTAAGAGATGTAAGAGATGTAAGAGATGTAAGAGAT 1560
Db 1501 TACGAAATGTAAGAGATGTAAGAGATGTAAGAGATGTAAGAGATGTAAGAGATGTAAGAGAT 1560
QY 1561 AGATTCACAAAGATGTAAGAGATGTAAGAGATGTAAGAGATGTAAGAGATGTAAGAGATGTA 1620
Db 1561 AGATTCACAAAGATGTAAGAGATGTAAGAGATGTAAGAGATGTAAGAGATGTAAGAGATGTA 1620
QY 1621 GCTTTCCACAGTTTGTATCTGCACTCACTCACTCACTCACTCACTCACTCACTCACTCACTCA 1680
Db 1621 GCTTTCCACAGTTTGTATCTGCACTCACTCACTCACTCACTCACTCACTCACTCACTCACTCA 1680
QY 1681 CAGAAATTAAGAAATTTAAGTCAAAATGTAAGATGTAAGATGTAAGATGTAAGATGTAAGAT 1740
Db 1681 CAGAAATTAAGAAATTTAAGTCAAAATGTAAGATGTAAGATGTAAGATGTAAGATGTAAGAT 1740
QY 1741 CTGCTGTCCGCTCAGCG 1800
Db 1741 CTGCTGTCCGCTCAGCG 1800
QY 1801 GCTGTATACGCCAACCCCTGTCTAGTGACTTAAGTGAAGCCGTCTGTGTGTGTGTGTGTGTGT 1860
Db 1801 GCTGTATACGCCAACCCCTGTCTAGTGACTTAAGTGAAGCCGTCTGTGTGTGTGTGTGTGTGT 1860
QY 1861 ATCCAGCATCAACAGGCTGT 1920
Db 1861 ATCCAGCATCAACAGGCTGT 1920
QY 1921 AGAGCAGAGATGTAAGT 1980
Db 1921 AGAGCAGAGATGTAAGT 1980
QY 1981 CTTTAAATGAAGAGTGAAGTTCCTCAACTTAA 2015
Db 1981 CTTTAAATGAAGAGTGAAGTTCCTCAACTTAA 2015

RESULT 2
AF272148 2015 bp mRNA linear PRI 21-JUN-2000
LOCUS AF272148 Homo sapiens KRAB zinc finger protein (RITA) mRNA, complete cds.
DEFINITION AF272148
ACCESSION AF272148
VERSION AF272148.1 GI:8575774
KEYWORDS
SOURCE Homo sapiens (human)

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominoidea; Homo.
REFERENCE 1 (bases 1 to 2015)
AUTHORS Ripe, V., Belge, G., Melboom, M., Kazmierczak, B., Fueco, A. and
Bullerdick, J.
TITLE A KRAB zinc finger protein gene is the potential target of 19q13
translocation in benign thyroid tumors
JOURNAL Genes Chromosomes Cancer 26 (3), 229-236 (1999)
PUBMED 10502321
2 (bases 1 to 2015)
AUTHORS Ripe, V., Melboom, M. and Bullerdick, J.
TITLE Direct Submission
JOURNAL Submitted (25-May-2000) Center for Human Genetics and Genetic
Counseling, University of Bremen, Loebener Strasse 2, Bremen 28359,
Germany

FEATURES
source Location/Qualifiers
1..2015
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="19"
/map="19q13.4"
1..2015
/gene="RITA"
/note="corresponds to exons 1 through 5"

gene
1..186
/gene="RITA"
/note="corresponds to exons 1 through 5"
5'UTR
1..186
/gene="RITA"
187..1578
/gene="RITA"
/note="Kruppel-type zinc finger protein"
/codon_start=1
/product="KRAB zinc finger protein"
/protein_id="AAF78075.1"
/db_xref="GI:8575775"

CDS
/translation="MAOGLVTFADVAIDFSOEMACINSABRDLYMDWLLENYSILVS
LDLESAYENKSLPTEKNITHETIAKSNRSRKSISGRNITCGETLERPORSKRYNO
MINTVVKPATREGETPRTTHQHHKENSFECDCKAFSRGQLSDHQKHTEKPYE
CECKAFRPMQQLTOHOKIHTGEKPYECKDCKAFRMSLSIVHKRIHTGEKPYE
DCKAFRMSDELTOHOKIHTGEKPYECKDCKAFRMSLSIVHKRIHTGEKPYE
GKAFICGSLIOHKRIHTGEKPYECKDCKAFRMSLSIVHKRIHTGEKPYE
AFRMSLSIVHKRIHTGEKPYECKDCKAFRMSLSIVHKRIHTGEKPYE
IYSSSLIVHKRIHTGEKPYECKDCKAFRMSLSIVHKRIHTGEKPYE
LNLHREHORHINS"
misc_feature
196..321
/gene="RITA"
/note="Region: KRAB A domain"
misc_feature
577..1575
/gene="RITA"
/note="Region: zinc finger domain"
3'UTR
1579..2015
/gene="RITA"
1685..1690
/gene="RITA"

polyA_signal
/gene="RITA"
ORIGIN

Query Match 100.0%; Score 2015; DB 8; Length 2015;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2015; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGCTTCTGACCGGTCGACGCAACCGTGTCTCCGCCAGTCCGCGCAGCCGACATCTT 60
Db 1 CGCTTCTGACCGGTCGACGCAACCGTGTCTCCGCCAGTCCGCGCAGCCGACATCTT 60
QY 61 TCAGAAAAAGCATCCCGCAGAGAGAAAGCAATCGTTAAACATCTTAGCTCAGCTTAC 120
Db 61 TCAGAAAAAGCATCCCGCAGAGAGAAAGCAATCGTTAAACATCTTAGCTCAGCTTAC 120
QY 121 CTCTGGAAATTTGCTTCTTCAAGTGAAAACCCGAGAAAGATGATCATGTTCTTCA 180
Db 121 CTCTGGAAATTTGCTTCTTCAAGTGAAAACCCGAGAAAGATGATCATGTTCTTCA 180

QY 121 CTCTGGAAATTTGCTTCTTCAAGTGAAAACCCGAGAAAGATGATCATGTTCTTCA 180

QY 181 AAAACAATGGCCAGAGGTTTGGTGAAGTTGCGCGAGCGTAGCCATAGACTTTTCTAGAG 240
Db 181 AAAACAATGGCCAGAGGTTTGGTGAAGTTGCGCGAGCGTAGCCATAGACTTTTCTAGAG 240
QY 241 GAGTGGGCTGTCTGAACCTGTGCTCAGAGGAGCCTGTACTGGGACGTGATGCTGAGAAC 300
Db 241 GAGTGGGCTGTCTGAACCTGTGCTCAGAGGAGCCTGTACTGGGACGTGATGCTGAGAAC 300
QY 301 TACAGTAATCTGTGCTCAGTGAATTTGGAGTACAGATATGAAAATTAAGATTACTTACA 360
Db 301 TACAGTAATCTGTGCTCAGTGAATTTGGAGTACAGATATGAAAATTAAGATTACTTACA 360
QY 361 GAAAAAATCATTCATGAATAAGAGCTTCCAAAAGAAATTCAGATAGAAAGTAATCC 420
Db 361 GAAAAAATCATTCATGAATAAGAGCTTCCAAAAGAAATTCAGATAGAAAGTAATCC 420
QY 421 CTGGCCGTAATCTGATATGTAAGAGTACGCTTGAAGACCAAGCGCTCAGAGGAGAG 480
Db 421 CTGGCCGTAATCTGATATGTAAGAGTACGCTTGAAGACCAAGCGCTCAGAGGAGAG 480
QY 481 TATGTCAATCAGATATCATCAATTTATGCAAAAGCCTGTACTAGAGAGAGCCCT 540
Db 481 TATGTCAATCAGATATCATCAATTTATGCAAAAGCCTGTACTAGAGAGAGCCCT 540
QY 541 CCTAGAACACATCAGAGCATCATTAAGAGAAATTCCTTGAATGTAAGGACTGTGGAG 600
Db 541 CCTAGAACACATCAGAGCATCATTAAGAGAAATTCCTTGAATGTAAGGACTGTGGAG 600
QY 601 GCCTTAACTGTCGCTATCAACTTATGCAACATCAGAAATTCATATGCTGAGAAACCT 660
Db 601 GCCTTAACTGTCGCTATCAACTTATGCAACATCAGAAATTCATATGCTGAGAAACCT 660
QY 661 TATGATGTAAGAAATGTAAGAGGCTTCGTTGGGCAATCAGCTTACATCAATCAA 720
Db 661 TATGATGTAAGAAATGTAAGAGGCTTCGTTGGGCAATCAGCTTACATCAATCAA 720
QY 721 AAAATTCATACGCGGAGAGAGCCCTAGAAATGTAAGACTGTGGAGAGGCTTTGATGG 780
Db 721 AAAATTCATACGCGGAGAGAGCCCTAGAAATGTAAGACTGTGGAGAGGCTTTGATGG 780
QY 781 GGCCTCAACCTCGTATTCATTAAGAGATTCATACGCTGTAAGAAACCTATGTAAGTAA 840
Db 781 GGCCTCAACCTCGTATTCATTAAGAGATTCATACGCTGTAAGAAACCTATGTAAGTAA 840
QY 841 GACTGTGAAAAGGCTTTGCGCTGTATGAGTCACTCAGACACAGAGATTCCACT 900
Db 841 GACTGTGAAAAGGCTTTGCGCTGTATGAGTCACTCAGACACAGAGATTCCACT 900
QY 901 GGGGAGAAAGCTAGAAATGCAAGACTGTGGAGAGCCTTAGCGGTGTATTAACCT 960
Db 901 GGGGAGAAAGCTAGAAATGCAAGACTGTGGAGAGCCTTAGCGGTGTATTAACCT 960
QY 961 ATTGACACAGAGAAATTCATTAAGTGGAGAGAGCCTTACGAGTGTAAAGACTGTGGAG 1020
Db 961 ATTGACACAGAGAAATTCATTAAGTGGAGAGAGCCTTACGAGTGTAAAGACTGTGGAG 1020
QY 1021 GCTTTAATTTGTGTTCAAGCTCATTCAGCATTAAGAAATTCACAGGTGAGAAACCC 1080
Db 1021 GCTTTAATTTGTGTTCAAGCTCATTCAGCATTAAGAAATTCACAGGTGAGAAACCC 1080
QY 1081 TATGAAATGTCAGAAATGAGAGGCTTTACTGAGTCAATTAACCTTACTGAGATCAG 1140
Db 1081 TATGAAATGTCAGAAATGAGAGGCTTTACTGAGTCAATTAACCTTACTGAGATCAG 1140
QY 1141 AAGATCCACACGCGTGAAGAGCTTCAGAAATGTAAGAGTGTGGAGAGGCTTTGCTGG 1200
Db 1141 AAGATCCACACGCGTGAAGAGCTTCAGAAATGTAAGAGTGTGGAGAGGCTTTGCTGG 1200
QY 1201 GGTTCGAGCTTCGTTAAGCAGAGAGTACATTAAGGCGAGAGAGCCTTCAAGTGCACA 1260
Db 1201 GGTTCGAGCTTCGTTAAGCAGAGAGTACATTAAGGCGAGAGAGCCTTCAAGTGCACA 1260
QY 1261 GAATGTGGAAGAGGCTTCAATTTGTTGATATCACTCAGACAGAGAAATTCACACA 1320

```

|||||
Db 1261 GAATGTGGAGAGCCCTTCAATTGTGGCTATCACTCAGTCAGACAGAGAAATCCACACA 1320
QY 1321 GGGCGAAACCCCGTATTAATGTAAGAGTGTTGGGAAGGCTTTCACTTAATGATCGACCTTC 1380
Db 1321 GGGCGAAACCCCGTATTAATGTAAGAGTGTTGGGAAGGCTTTCACTTAATGATCGACCTTC 1380
QY 1381 GTGAAACATGAGAGAAATTCATACCGGGGTGAAACCTTATGGGTGTACAGAAATGTGGGAAG 1440
Db 1381 GTGAAACATGAGAGAAATTCATACCGGGGTGAAACCTTATGGGTGTACAGAAATGTGGGAAG 1440
QY 1441 AGCTTTAGTCAGGCGCATAGCTTACACAAATCAGAAAACGCAAGTGGGCGAAATCC 1500
Db 1441 AGCTTTAGTCAGGCGCATAGCTTACACAAATCAGAAAACGCAAGTGGGCGAAATCC 1500
QY 1501 TACGAATGTAAAGAGTGGGGAAGGATGTAAACCTTAAACCATCTCCGAGAACATCAG 1560
Db 1501 TACGAATGTAAAGAGTGGGGAAGGATGTAAACCTTAAACCATCTCCGAGAACATCAG 1560
QY 1561 AGGATCCACAACAGTTGAAAGAGCCTTTTGAACGAGTAGCCGCTGATCTATGTTTC 1620
Db 1561 AGGATCCACAACAGTTGAAAGAGCCTTTTGAACGAGTAGCCGCTGATCTATGTTTC 1620
QY 1621 GCTTTTCACAGTTTGTACTCTGAGTCAACTGCACTTCAAAAATATTAAATGAAATTC 1680
Db 1621 GCTTTTCACAGTTTGTACTCTGAGTCAACTGCACTTCAAAAATATTAAATGAAATTC 1680
QY 1681 CAGAAATTAAGAAATTTTAAGTCTCAATAGTGTTGCCCTTGTGAGTGGCGGATGAATCT 1740
Db 1681 CAGAAATTAAGAAATTTTAAGTCTCAATAGTGTTGCCCTTGTGAGTGGCGGATGAATCT 1740
QY 1741 CTGCTGTCCGCTCCAGCCGCGGAGATGTAGTCAATCCCTTGTGTCAGACATCCAC 1800
Db 1741 CTGCTGTCCGCTCCAGCCGCGGAGATGTAGTCAATCCCTTGTGTCAGACATCCAC 1800
QY 1801 GCTGTATAGCCCAACCCCTGCTAGTGAATTTAGTAGCCCTTGTGATCAATCAACT 1860
Db 1801 GCTGTATAGCCCAACCCCTGCTAGTGAATTTAGTAGCCCTTGTGATCAATCAACT 1860
QY 1861 ATCCCAAGCATCAAGTGCCTGTGCCCAAGTAGTCTTCACTTGTAAACAGTGGCCCAAG 1920
Db 1861 ATCCCAAGCATCAAGTGCCTGTGCCCAAGTAGTCTTCACTTGTAAACAGTGGCCCAAG 1920
QY 1921 AGAGCAGAGTAGTAGTGTCTGTGATTCGATATGCCAAAGAGAGCCAAAGTGTCTC 1980
Db 1921 AGAGCAGAGTAGTAGTGTCTGTGATTCGATATGCCAAAGAGAGCCAAAGTGTCTC 1980
QY 1981 CTTTAAATGAAAGGTGAAAGTTCTCACTTAA 2015
Db 1981 CTTTAAATGAAAGGTGAAAGTTCTCACTTAA 2015

RESULT 3
AX427523 2098 bp DNA linear PAT 20-JUN-2002
LOCUS AX427523
DEFINITION Sequence 33 from Patent WO0121836.
ACCESSION AX427523
VERSION AX427523.1 GI:21537669
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
            Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Homiidae; Homo.
REFERENCE
1 Hodgson, D.M., Lincoln, S.E., Russo, F.D., Spiro, P.A., Banville, S.C.,
  Bratcher, Shaw, R., Dufour, G.F., Cohen, H.U., Rosen, B.H., Shah, P.,
  Chalup, M.S., Hillman, J.L., Jones, A.L., Yu, J.Y., Greenawald, L.B.,
  Panzer, S.R., Roseberry, A.M., Wright, R.J., Chen, W., Liu, T.,
  Yap, P.E., Stockbreher, T.K., Ameshey, S. and Fong, W.T.
  Molecules for diagnostics and therapeutics
  Patent: WO 0121836-A 33 29-MAR-2001;
  Incyte Genomics, Inc. (US)

```

```

FEATURES
    source
        location/Qualifiers
            1..2098
            /organism="Homo sapiens"
            /mol_type="unassigned DNA"
            /db_xref="taxon:9606"
            /note="Incyte ID No: 902791.3.dec"
ORIGIN
Query Match      98.6%; Score 1986.8; DB 6; Length 2098;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 2010; Conservative 0; Mismatches 2; Indels 2; Gaps 2;

QY 1 CGCTTCTGCACCGGTGAGCAACCGTGTCTTCCGCCAGTCCGCGCAGGCGCAATCCT 60
Db 85 CGCTTCTGCACCGGTGAGCAACCGTGTCTTCCGCCAGTCCGCGCAGGCGCAATCCT 144
QY 61 TCAGAAAAAGCATCCCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 120
Db 145 TCAGAAAAAGCATCCCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 204
QY 121 CTCTGGAAATTTGTCTTCTCAAGTGAACCCCGAGAGAGAGAGAGAGAGAGAGAGAG 180
Db 205 CTCTGGAAATTTGTCTTCTCAAGTGAACCCCGAGAGAGAGAGAGAGAGAGAGAGAG 264
QY 181 AAAACATAGGCGCGAGGTGTGAGCTTCGCGAGAGAGAGAGAGAGAGAGAGAGAGAG 240
Db 265 AAAACATAGGCGCGAGGTGTGAGCTTCGCGAGAGAGAGAGAGAGAGAGAGAGAGAG 324
QY 241 GAGTGGGCGCTGTCTGAACCTGTCTCAGAGGAGCTGTCTAGGAGAGAGAGAGAGAGAG 300
Db 325 GAGTGGGCGCTGTCTGAACCTGTCTCAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAG 384
QY 301 TACAGTAATCTGATCTCACTGAGATTTGAGTCAAGTCAATGAAATTAAGATTACTTACA 360
Db 385 TACAGTAATCTGATCTCACTGAGATTTGAGTCAAGTCAATGAAATTAAGATTACTTACA 444
QY 361 GAAAAAACAATTCATGAATTAAGAGGCTTCCAAAGAGAGAGAGAGAGAGAGAGAGAGAG 420
Db 445 GAAAAAACAATTCATGAATTAAGAGGCTTCCAAAGAGAGAGAGAGAGAGAGAGAGAGAG 504
QY 421 CTGGGCGGTAACTGATATGTGAAGGTGAGCTTGAAGAGAGAGAGAGAGAGAGAGAGAG 480
Db 505 CTGGGCGGTAACTGATATGTGAAGGTGAGCTTGAAGAGAGAGAGAGAGAGAGAGAGAG 564
QY 481 TATGTCAATCAAGATGATCAATTAATGTCAAAAGGCTGTCTAAGAGAGAGAGAGAGAG 539
Db 565 TATGTCAATCAAGATGATCAATTAATGTCAAAAGGCTGTCTAAGAGAGAGAGAGAGAG 624
QY 540 TCTTAGAACATCAGAGACATTAAGAGAGATTCCTTGAATGTGAAGAGAGAGAGAGAGAG 599
Db 625 TCTTAGAACATCAGAGACATTAAGAGAGATTCCTTGAATGTGAAGAGAGAGAGAGAGAG 684
QY 600 GGCCTTTAGTCGAGGTATCAATCTTGAATCAATCAATCAATCAATCAATCAATCAATCA 659
Db 685 GGCCTTTAGTCGAGGTATCAATCTTGAATCAATCAATCAATCAATCAATCAATCAATCA 744
QY 660 TTATGATGTAAAGATGTAAAGAGGCTTCCGTTGGGCAATCACTTCACTCAATCA 719
Db 745 TTATGATGTAAAGATGTAAAGAGGCTTCCGTTGGGCAATCACTTCACTCAATCA 804
QY 720 AAAAATTCATCTCGGAGAGAGGCTTCCGTTGGGCAATCACTTCACTCAATCA 779
Db 805 AAAAATTCATCTCGGAGAGAGGCTTCCGTTGGGCAATCACTTCACTCAATCA 864
QY 780 GGGCTCAGCGCTGTATTCATTAAGAGAGATCACTTCACTTCACTTCACTTCACTTCA 838
Db 865 GGGCTCAGCGCTGTATTCATTAAGAGAGATTCATCTTCACTTCACTTCACTTCACTTCA 924
QY 839 AAGACTGTGAAAGGCTTCCGCGTGTGATAGTCACTCAGACACAGAGATTCACACA 898
Db 925 AAGACTGTGAAAGGCTTCCGCGTGTGATAGTCACTCAGACACAGAGATTCACACA 984
QY 899 CTGGGAGAAAGACTACGAATGCAAAAGCTGTGGAGAGAGAGAGAGAGAGAGAGAGAGAG 958

```

Db 985 CTGGGGAGAAAACCTACGAATGCAGAACTGTGGAGAGCTTTAGCCCTGTGTATAAC 1044
Qy 959 TTATTCAGACAAAGAAATTCATAGTGGGAGAAAGCTTTACGAGTGTAAAGACTGTGGGA 1018
Db 1045 TTATTCAGACAAAGAAATTCATAGTGGGAGAAAGCTTTACGAGTGTAAAGACTGTGGGA 1104
Qy 1019 AGGCTTTATTTGTGTGTCAACCTCATTCACGATTTAAAGAAATTCACACAGGTGAGAAAC 1078
Db 1105 AGGCTTTATTTGTGTGTCAACCTCATTCACGATTTAAAGAAATTCACACAGGTGAGAAAC 1164
Qy 1079 CCTATGAATGTCAAGAAATGTGGGAGAGCCCTTACTGAGTCAATTAACCTTACTCAGCATC 1138
Db 1165 CCTATGAATGTCAAGAAATGTGGGAGAGCCCTTACTGAGTCAATTAACCTTACTCAGCATC 1224
Qy 1139 AGAAGATCCACACCGGTGAGAAAGCTCAGCAATGTAAAGAGTGTGGGAGAGCCCTTTCCT 1198
Db 1225 AGAAGATCCACACCGGTGAGAAAGCTCAGCAATGTAAAGAGTGTGGGAGAGCCCTTTCCT 1284
Qy 1139 GGGGCTTCGAGCTCGTTAAGCAGAGAGATCATACGGGGCGAGAAAGCCGTACAAGTGA 1258
Db 1285 GGGGCTTCGAGCTCGTTAAGCAGAGAGATCATACGGGGCGAGAAAGCCGTACAAGTGA 1344
Qy 1259 CAGAATGTGGGAGAGCCCTCAATGTGTCTATCACTCAGCAGAGAGATCCACA 1318
Db 1345 CAGAATGTGGGAGAGCCCTCAATGTGTCTATCACTCAGCAGAGAGATCCACA 1404
Qy 1319 CAGGCGAAAACCCGTAATAATGTAAAGAGTGTGGGAGAGCCCTTCAATTAATGATCGAGCC 1378
Db 1405 CAGGCGAAAACCCGTAATAATGTAAAGAGTGTGGGAGAGCCCTTCAATTAATGATCGAGCC 1464
Qy 1379 TGTGAAACATGAGAAATTCATACGGGGGTGAAAACCTTATGGGTGTACAGATGTGGGA 1438
Db 1465 TGTGAAACATGAGAAATTCATACGGGGGTGAAAACCTTATGGGTGTACAGATGTGGGA 1524
Qy 1439 AGAGCTTTAGTACGCGGCATCAGCTTACACAAATCAGAAAACGACAGTGGGGCGAAAT 1498
Db 1525 AGAGCTTTAGTACGCGGCATCAGCTTACACAAATCAGAAAACGACAGTGGGGCGAAAT 1584
Qy 1499 CCTACGAATGTAAAGAGTGTGGGAGAGCATGTAAACCACTTAAACCATCTCCGAGAACATC 1558
Db 1585 CCTACGAATGTAAAGAGTGTGGGAGAGCATGTAAACCACTTAAACCATCTCCGAGAACATC 1644
Qy 1559 AGAGGATCCACAAAGCTTGAAGAGCTTTTGAACCGAGTACCCGCTGTATCTATGCTT 1618
Db 1645 AGAGGATCCACAAAGCTTGAAGAGCTTTTGAACCGAGTACCCGCTGTATCTATGCTT 1704
Qy 1619 TGGCTTTCCACAGTTGTACCGGAGTCACTGACAGTCAAAAATATTAAATGGAAAT 1678
Db 1705 TGGCTTTCCACAGTTGTACCGGAGTCACTGACAGTCAAAAATATTAAATGGAAAT 1764
Qy 1679 TCCAGAAATAAAGAAATTTTAAGTCTCAAAATGTGTGCTCTCTGAGTAGCGTAGTAAT 1738
Db 1765 TCCAGAAATAAAGAAATTTTAAGTCTCAAAATGTGTGCTCTCTGAGTAGCGTAGTAAT 1824
Qy 1739 CTCTGCTGTCCGGCTCCAGCCGCCGGGAGATGTGATCATCCCTTGTGCTCAGCATCC 1798
Db 1825 CTCTGCTGTCCGGCTCCAGCCGCCGGGAGATGTGATCATCCCTTGTGCTCAGCATCC 1884
Qy 1799 AGGCTGTATAGGCAACCCGCTGCTAGTACTTGTAGCGCTTGTGGTGCAGATCA 1858
Db 1885 AGGCTGTATAGGCAACCCGCTGCTAGTACTTGTAGCGCTTGTGGTGCAGATCA 1944
Qy 1859 CTATCCAGCATCAGTGTCTGTGCTGCAAGTATGCTTCACTTGTGTTAAAGTGGCCCC 1918
Db 1945 CTATCCAGCATCAGTGTCTGTGCTGCAAGTATGCTTCACTTGTGTTAAAGTGGCCCC 2004
Qy 1919 AGAGAGCAGAGATGATGTGTGTGTATCGATAGCCAAAGAGAAAGCACAAGTGTGT 1978
Db 2005 AGAGAGCAGAGATGATGTGTGTGTATTCGATATGCAAGAGAGAAAGCACAAGTGTGT 2064
Qy 1979 TCCTTTTAATGAAAAAGTGAAGTTCTCAACTT 2012
|||||

Db 2065 TCCTTTTAATGAAAAAGTGAAGTTCTCAACTT 2098
RESULT 4
AX274843 2173 bp DNA linear PAT 29-OCT-2001
LOCUS Sequence 108 from Patent WO0172777.
DEFINITION
ACCESSION AX274843
VERSION AX274843.1 GI:16547475
KEYWORDS
SOURCE
ORGANISM
Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.
REFERENCE
1 Hillman, J.L., Baughn, M.R., Yue, H., Lal, P., Lu, D.A., Patterson, C.,
Azimzai, Y., Bandman, O., Tang, Y.T., Mathur, P., Shah, P., Au-Young, J.,
and Reddy, R.
TITLE Transcription factors
JOURNAL Patent: WO 0172777-A 108 04-OCT-2001;
Incyte Genomics, Inc. (US)
FEATURES
source
1. 2173
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
/note="Incyte ID No: 095210CB1"
ORIGIN
Query Match 97.6%; Score 1967.4; DB 6; Length 2173;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1968; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 47 CAGGCGACATCTCTCAGAAAAAGCATCCCGAGGAGAGACGATGTTAAACATCTT 106
Db 189 CAGGCGACATCTCTCAGAAAAAGCATCCGAGGAGAGACGATGTTAAACATCTT 248
Qy 107 AGGTCAAGCTTACGCTCTCGGAATTTGTTCTTCTTCAATGTGAAACCCCGAAGACTGATC 166
Db 249 AGGTCAAGCTTACGCTCTCGGAATTTGTTCTTCTTCAATGTGAAACCCCGAAGACTGATC 308
Qy 167 AGTCTTCAAGTTCTTAAACAAATGCGCCAGGGTTGTGTAACGTTGCGACGTAAGCATAG 226
Db 309 AGTCTTCAAGTTCTTAAACAAATGCGCCAGGGTTGTGTAACGTTGCGACGTAAGCATAG 368
Qy 227 ACTTTCTCAGAGAGAGTGGGCTGTGTGAACCTGTCTCAGAGGAGCTGTACTGGAGCG 286
Db 369 ACTTTCTCAGAGAGAGTGGGCTGTGTGTGAACCTGTCTCAGAGGAGCTGTACTGGAGCG 428
Qy 287 TGATGCTGAGAACTACAGTAACTTGTGTCTCACTGAAATTTGGAGTCAACATATGAAATA 346
Db 429 TGATGCTGAGAACTACAGTAACTTGTGTCTCACTGAAATTTGGAGTCAACATATGAAATA 488
Qy 347 AGAGTTTACCTACAGAAAAAAACATTCATGAAATTAAGGCTTCCAAAAGAAATTCAGATA 406
Db 489 AGAGTTTACCTACAGAAAAAAACATTCATGAAATTAAGGCTTCCAAAAGAAATTCAGATA 548
Qy 407 GAAGAAGTAAATCCCTTGGCCGTAACTGATATGTGAAGTATCGCTTGAAGAACCACAGC 466
Db 549 GAAGAAGTAAATCCCTTGGCCGTAACTGATATGTGAAGTATCGCTTGAAGAACCACAGC 608
Qy 467 GCTCCAGAGGAGATATGCTCAATCAGATGATCATCAATTAATGTCAAAAAGCCCTGCTACTA 526
Db 609 GCTCCAGAGGAGATATGCTCAATCAGATGATCATCAATTAATGTCAAAAAGCCCTGCTACTA 668
Qy 527 GAGAAGGACCCCTCTTGAACACATCAGAGACATCAAGAGAAATTCCTTGAATGTA 586
Db 669 GAGAAGGACCCCTCTTGAACACATCAGAGACATCAAGAGAAATTCCTTGAATGTA 728
Qy 587 AGAGCTGTGGAAAGCCCTTATGCTGTGCTATCAACTTAATGTCATCAGAAAAATTCATA 646
Db 729 AGAGCTGTGGAAAGCCCTTATGCTGTGCTATCAACTTAATGTCATCAGAAAAATTCATA 788
|||||

```

QY 647 CTGGTGAAGAACTTATGAAATGTAAGAAATGTAAGAGGCTTCCGTTGGGGCAATCAGC 706
DB 789 CTGGTGAAGAACTTATGAAATGTAAGAAATGTAAGAGGCTTCCGTTGGGGCAATCAGC 848
QY 707 TTACTCAACATCAAAAAATTCATCTGGGGAGAGGCTTCAGATGTAAAGACTGGGA 766
DB 849 TTACTCAACATCAAAAAATTCATCTGGGGAGAGGCTTCAGATGTAAAGACTGGGA 908
QY 767 AGGCTTTGAGATGGGGCTCAAGGCTGTTATTCATTAAGAGATTCATCTGGTGAAGAAC 826
DB 909 AGGCTTTGAGATGGGGCTCAAGGCTGTTATTCATTAAGAGATTCATCTGGTGAAGAAC 968
QY 827 CCTATGAATGTAAAGACTGTGAGAAAGGCTTTGCGGTGTGTGAAGAGCTCACTCAGCAC 886
DB 969 CCTATGAATGTAAAGACTGTGAGAAAGGCTTTGCGGTGTGTGAAGAGCTCACTCAGCAC 1028
QY 887 AAGAAATTCACACCTGGGGAGAAAGACTAGAAATGCAAAAGCTGTGGGAGAGACTTTAGCC 946
DB 1029 AAGAAATTCACACCTGGGGAGAAAGACTAGAAATGCAAAAGCTGTGGGAGAGACTTTAGCC 1088
QY 947 GTGTGTATTAATCTTATTCAGACACAGAAATTCATAGTGGGGAGAGGCTTACAGTGT 1006
DB 1089 GTGTGTATTAATCTTATTCAGACACAGAAATTCATAGTGGGGAGAGGCTTACAGTGT 1148
QY 1007 AAGACTGTGGGAAGGCTTTTATTTGTGTCAAGGCTCATTCAGCATAAAGAAATTCACA 1066
DB 1149 AAGACTGTGGGAAGGCTTTTATTTGTGTCAAGGCTCATTCAGCATAAAGAAATTCACA 1208
QY 1067 CAGGTGAAGAAACCTTATGAAATGTCAGAAATGTGGGAAGGCTTTATCTCGAGTCAATTA 1126
DB 1209 CAGGTGAAGAAACCTTATGAAATGTCAGAAATGTGGGAAGGCTTTATCTCGAGTCAATTA 1268
QY 1127 TTACTCAGCATCAGAAATCACAACCGGTGAGAAAGGCTTCAGATGTAAAGAGTGTGGGA 1186
DB 1269 TTACTCAGCATCAGAAATCACAACCGGTGAGAAAGGCTTCAGATGTAAAGAGTGTGGGA 1328
QY 1187 AGGCTTTGCGTGGGGTTCGAGGCTCGTTAAGCAGCAGAGATATCATCCGGGCGAGAGAG 1246
DB 1329 AGGCTTTGCGTGGGGTTCGAGGCTCGTTAAGCAGCAGAGATATCATCCGGGCGAGAGAG 1388
QY 1247 CGTACAGATGTCAGAAATGTGGGAAGGCTTCATTTGTGGCTATCACTTCACTCAGACG 1306
DB 1389 CGTACAGATGTCAGAAATGTGGGAAGGCTTCATTTGTGGCTATCACTTCACTCAGACG 1448
QY 1307 AAGAAATTCACACAGGCGAAACCCGATATTAATGTAAAGAGTGTGGGAAGGCTTTCA 1366
DB 1449 AAGAAATTCACACAGGCGAAACCCGATATTAATGTAAAGAGTGTGGGAAGGCTTTCA 1508
QY 1367 ATGATTCGAGGCTCGTGAACCATGAGAAATTCATACCGGGGTGAACCCCTATGGGTGA 1426
DB 1509 ATGATTCGAGGCTCGTGAACCATGAGAAATTCATACCGGGGTGAACCCCTATGGGTGA 1568
QY 1427 CAGAAATGTGGGAAGAGCTTTATGTCACGGCCATCAGCTTACACAAATCAAAAAACGACA 1486
DB 1569 CAGAAATGTGGGAAGAGCTTTATGTCACGGCCATCAGCTTACACAAATCAAAAAACGACA 1628
QY 1487 GTGGGGCGAAATCTCTAGAAATGTAAAGAGTGGGGGAAGGAGATGTAAACCTTAAACCATC 1546
DB 1629 GTGGGGCGAAATCTCTAGAAATGTAAAGAGTGGGGGAAGGAGATGTAAACCTTAAACCATC 1688
QY 1689 TCGGAGAAATCAAGAGATCAACAACAGTTGAAGAGCTTTTGAAGCAGATGACCGGCTC 1748
DB 1607 GTATCTATGTGTTTGGCTTTCCACAGTTGTTTACTGTGACATCAGAGTTCAAAATAT 1666
QY 1749 GTATCTATGTGTTTGGCTTTCCACAGTTGTTTACTGTGACATCAGAGTTCAAAATAT 1808
DB 1667 TAAATGGAATAATCCAGAAATTAAGAAATTTTAAGTCAATAGGTGGGCTTCTGAGTA 1726
QY 1809 TAAATGGAATAATCCAGAAATTAAGAAATTTTAAGTCAATAGGTGGGCTTCTGAGTA 1868

```

```

QY 1727 GCGTGAATAATCTCTCGCTGTCCGGCTTCAGCCGGCGGGGATGTGAATCATCCCTTGG 1786
DB 1869 GCGTGAATAATCTCTCGCTGTCCGGCTTCAGCCGGCGGGGATGTGAATCATCCCTTGG 1928
QY 1787 TCCAGACATTCACGCTGTATTCGACACCCACCTGTAGTACCTTAGAGCGGTCTTG 1846
DB 1929 TCCAGACATTCACGCTGTATTCGACACCCACCTGTAGTACCTTAGAGCGGTCTTG 1988
QY 1847 TGATCAGATCAACTATCCAGCATCAGTGTGCTGTGCCAAGTAGTCTCACTTTGCTT 1906
DB 1989 TGATCAGATCAACTATCCAGCATCAGTGTGCTGTGCCAAGTAGTCTCACTTTGCTT 2048
QY 1907 AACAGTGGCCCCAGAGAGCAGAGATGTGATGTGTGTGATTTGGATATGCCAAGGAAG 1966
DB 2049 AACAGTGGCCCCAGAGAGCAGAGATGTGATGTGTGTGATTTGGATATGCCAAGGAAG 2108
QY 1967 CCACAAAGTCTCTCTTTAAATGAAAGGTAAGATTCTCAACTTAA 2015
DB 2109 CCACAAAGTCTCTCTTTAAATGAAAGGTAAGATTCTCAACTTAA 2157

```

```

RESULT 5
LOCUS BD127591 2169 bp DNA linear PART 18-SRP-2002
DEFINITION Primer for synthesizing full-length cDNA and use thereof.
ACCESSION BD127591
VERSION BD127591.1 GI:23222536
KEYWORDS JP 2002017375-A/3022.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

```

```

REFERENCE
AUTHORS Ota,T., Nishikawa,T., Isogai,T., Hayashi,K., Ishii,S., Kawai,Y.,
Wakamatsu,A., Sugiyama,T., Nagai,K., Kojima,S., Otsuki,T. and
Homindae,Homo.
1 (bases 1 to 2169)
Ota,T., Nishikawa,T., Isogai,T., Hayashi,K., Ishii,S., Kawai,Y.,
Wakamatsu,A., Sugiyama,T., Nagai,K., Kojima,S., Otsuki,T. and
Koga,H.
Primer for synthesizing full-length cDNA and use thereof
Patent: JP 2002017375-A 3022 22-JAN-2002;
HELIX RESEARCH INSTITUTE
OS Homo sapiens (human)
PN JP 2002017375-A/3022
PE 22-JAN-2002
PF 07-JUL-2000 JP 2000253172
PI TOSHIO OTA,TETSUO NISHIKAWA,TAKAO ISOGAI,KOJI HAYASHI,SHIZUKO
PI ISHII,
PI YURI KAWAI,AI WAKAMATSU,TOMOYASU SUGIYAMA,KEIICHI MAGAI, PI
SHINICHI KOJIMA,
PI TETSUOI OTSUKI,HISASHI KOGA
PC
C12N15/09,C07K14/47,C07K16/18,C12N1/15,C12N1/19,C12N1/21,C12N5/ PC
10, C12P21/02,C12Q1/68//C12P21/08,G06F17/30,C12N15/00,C12N5/00 CC
Primer for synthesizing full-length cDNA and use thereof FH key
Location/Qualifiers
(344)..(1540).
FT CDS

```

```

FEATURES
source 1..2169
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
ORIGIN

```

```

Query Match 97.2%; Score 1958; DB 6; Length 2169;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 1961; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

```

```

QY 47 CAGGCCAGATCTTTCAGAAAAAGCATCCCGAGAGAGAGAGAAATCGTTAAACATCTT 106
DB 204 CAGGCCAGATCTTTCAGAAAAAGCATCCCGAGAGAGAGAGAAATCGTTAAACATCTT 263
QY 107 AGGTCAAGCTTACCTCTCGGAATTTGTCTTCTTCAGTGAAGAACCCCGAAGAACTGATC 166

```


Db 264 AGTTCAGCTAGCCTTCGGAATTTGTTCTTCTGAGTGAAGAACCCCGAGAAGACTGATC 323
 QY 167 AGTTCAGCTTAAACAATGCGCCAGGGTTTGTAACCTTCCCGACGTAAGCCATAG 226
 Db 324 AGTTCAGCTTAAACAATGCGCCAGGGTTTGTAACCTTCCCGACGTAAGCCATAG 383
 QY 227 ACTTTTCAGAGAGAGTGGGCTGTGAACTGTGCTCAGAGGAGCCTGTCTGGGACG 286
 Db 384 ACTTTTCAGAGAGAGTGGGCTGTGAACTGTGCTCAGAGGAGCCTGTCTGGGACG 443
 QY 287 TGATGCTGAGAACTACAGTAATTGATCTCACTGGAATTTGGAGTCAGCATATGAATA 346
 Db 444 TGATGCTGAGAACTACAGTAATTGATCTCACTGGAATTTGGAGTCAGCATATGAATA 503
 QY 347 AGAGTTTACCTACAGAAAAAAACATTCATGAATAAGGCTTCCAAAAAGGAATTGAGTA 406
 Db 504 AGAGTTTACCTACAGAAAAAAACATTCATGAATAAGGCTTCCAAAAAGGAATTGAGTA 563
 QY 407 GAAAGATTAATCCCTTGCGGTACCTGGAATATGTAAGGTACGCTTGAAAGACACAGC 466
 Db 564 GAAAGATTAATCCCTTGCGGTACCTGGAATATGTAAGGTACGCTTGAAAGACACAGC 623
 QY 467 GCTCCAGAGGAGTATGTCATCAGATGATCATCATTAATGTCAAAAAGCCTGCTACTA 526
 Db 624 GCTCCAGAGGAGTATGTCATCAGATGATCATCATTAATGTCAAAAAGCCTGCTACTA 683
 QY 527 GAGAAAGGCACTCTCTGAGACATCAAGACATCATTAAGGAAATTCCTTGAATGTA 586
 Db 684 GAGAAAGGCACTCTCTGAGACATCAAGACATCATTAAGGAAATTCCTTGAATGTA 743
 QY 587 AGGATGAGGGAAGGCTTTAGTGGTGGTATCACTTGAATCAATGAGAAATTCATTA 646
 Db 744 AGGATGAGGGAAGGCTTTAGTGGTGGTATCACTTGAATCAATGAGAAATTCATTA 803
 QY 647 CTGTGAGAAACCTTTGATGATGAATAAGATGTAAGAAGGCTTCCGTTGGGCAATCA 706
 Db 804 CTGTGAGAAACCTTTGATGATGAATAAGATGTAAGAAGGCTTCCGTTGGGCAATCA 863
 QY 707 TTACTCAACATCAAAAAAATTCATCTGGGAGAGAACCTTCAAGATGTAAGACTGTGGA 766
 Db 864 TTACTCAACATCAAAAAAATTCATCTGGGAGAGAACCTTCAAGATGTAAGACTGTGGA 923
 QY 767 AGGCTTTTCAGTGGGCTCAAGCCCTGTTATCATTAAGAGATGTAATCTGGTGAAGAAC 826
 Db 924 AGGCTTTTCAGTGGGCTCAAGCCCTGTTATCATTAAGAGATGTAATCTGGTGAAGAAC 983
 QY 827 CCTATGAATGTAAGAAGCTGTGAAAGGCTTTCGCGTGTGATGAGCTCACTCAGAC 886
 Db 984 CCTATGAATGTAAGAAGCTGTGAAAGGCTTTCGCGTGTGATGAGCTCACTCAGAC 1043
 QY 887 AGAGATTCACACTGGGAGAAAGACTACGAATGCAAGAGCTGTGGAAAGCCTTTAGCC 946
 Db 1044 AGAGATTCACACTGGGAGAAAGACTACGAATGCAAGAGCTGTGGAAAGCCTTTAGCC 1103
 QY 947 GGTGTATTAACCTTATTCAGCAAGAGAAATTCATTAAGGAGGAAAGCCTTACGAGTGA 1006
 Db 1104 GGTGTATTAACCTTATTCAGCAAGAGAAATTCATTAAGGAGGAAAGCCTTACGAGTGA 1163
 QY 1007 AAGACTGTGGGAAGGCTTTATTTGTGTTCAAGCTCATTCAGCATTAAGAAATTCACA 1066
 Db 1164 AAGACTGTGGGAAGGCTTTATTTGTGTTCAAGCTCATTCAGCATTAAGAAATTCACA 1223
 QY 1067 CAGGTGAGAAACCTTATGATGTCAGAAATGTGGAAAGGCTTTACTGAGTCAATTACC 1126
 Db 1224 CAGGTGAGAAACCTTATGATGTCAGAAATGTGGAAAGGCTTTACTGAGTCAATTACC 1283
 QY 1127 TTACTCAGCATCAGAAAGATCCACCGGTGAGAGCCTCAGAAATGTAAGAGTGTGGGA 1186
 Db 1284 TTACTCAGCATCAGAAAGATCCACCGGTGAGAGCCTCAGAAATGTAAGAGTGTGGGA 1343
 QY 1187 AGGCTTTTCGCTGGGTTTCAAGCTCTGTTAAGCAGAGAGTACATTCGCGCGAAGAC 1246
 Db 1344 AGGCTTTTCGCTGGGTTTCAAGCTCTGTTAAGCAGAGAGTACATTCGCGCGAAGAC 1403

QY 1247 GGTACAAATGACAGAAATGTGGAAAGGCTTCAATTTGGCTATCACTCAGACG 1306
 Db 1404 GGTACAAATGACAGAAATGTGGAAAGGCTTCAATTTGGCTATCACTCAGACG 1463
 QY 1307 AGAGATTCACACAGGCGCAACCCGTAATAATGTAAGAGTGTGGAAAGCTTTCATT 1366
 Db 1464 AGAGATTCACACAGGCGCAACCCGTAATAATGTAAGAGTGTGGAAAGCTTTCATT 1523
 QY 1367 ATGATCGAGCTTCTGTAACATGAGAAATTCATACCGGGGTGAACCTTATGGGTGA 1426
 Db 1524 ATGATCGAGCTTCTGTAACATGAGAAATTCATACCGGGGTGAACCTTATGGGTGA 1583
 QY 1427 CAGAAATGAGGAAGGCTTTATGTCACGGCCTACGCTTACCAATCAGAAAGCA 1486
 Db 1584 CAGAAATGAGGAAGGCTTTATGTCACGGCCTACGCTTACCAATCAGAAAGCA 1643
 QY 1487 GTGGGCGCAAAATCTTACAGAAATGTAAGAGTGTGGAAAGGCAATGTAACCACTTAAACATC 1546
 Db 1644 GTGGGCGCAAAATCTTACAGAAATGTAAGAGTGTGGAAAGGCAATGTAACCACTTAAACATC 1703
 QY 1547 TCCGAAACATCAGAGATCCACAAACAGTGAAGAGCTTTTGAACGAGTACCGCTC 1606
 Db 1704 TCCGAAACATCAGAGATCCACAAACAGTGAAGAGCTTTTGAACGAGTACCGCTC 1763
 QY 1607 GTATCTATGTTTGGCTTTTCCACAGTTTGTATCCTGCAATGCAATGCAAGTCAAAAAATAT 1666
 Db 1764 GTATCTATGTTTGGCTTTTCCACAGTTTGTATCCTGCAATGCAATGCAAGTCAAAAAATAT 1823
 QY 1667 TAAATGCAAAATTCAGAAATTAAGAAATTTTAAGTCTGCAAAATGTGTGCTCTGAGTA 1726
 Db 1824 TAAATGCAAAATTCAGAAATTAAGAAATTTTAAGTCTGCAAAATGTGTGCTCTGAGTA 1883
 QY 1727 GCGTGTAGAAATCTCTGCTGTCCGCTCCAGCGGCGGGAGTGAATCCCTTGG 1786
 Db 1884 GCGTGTAGAAATCTCTGCTGTCCGCTCCAGCGGCGGGAGTGAATCCCTTGG 1943
 QY 1787 TCCAGCAATCCAGCGCTTATACGCCAACCCCTGTAAGTACTTAAGCCGCTTGG 1846
 Db 1944 TCCAGCAATCCAGCGCTTATACGCCAACCCCTGTAAGTACTTAAGCCGCTTGG 2003
 QY 1847 TGATCAGATCAATTCACAGATCAAGTGTGTGCTGCTCAAGTGTGCTCACTTGGCTT 1906
 Db 2004 TGATCAGATCAATTCACAGATCAAGTGTGTGCTGCTCAAGTGTGCTCACTTGGCTT 2063
 QY 1907 AACAGTGGCCCCAGAGACAGAGAGTGAATGATGCTGTGATTCGATATGCCAAGAGAG 1966
 Db 2064 AACAGTGGCCCCAGAGACAGAGAGTGAATGATGCTGTGATTCGATATGCCAAGAGAG 2123
 QY 1967 CCACAAAAGTCTTCTTTTAAATGAAAAAGTGGAAGTCTCAACTT 2012
 Db 2124 CCACAAAAGTCTTCTTTTAAATGAAAAAGTGGAAGTCTCAACTT 2169

RESULT 6
 CQ783329
 LOCUS DEFINITION Sequence 3469 from Patent EP1396543.
 ACCESSION CQ783329
 VERSION CQ783329.1 GI:45503259
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Homidae; Homo.

REFERENCE
 AUTHORS
 1 Ota, T., Nishikawa, T., Isogai, T., Hayashi, K., Ishii, S., Kawai, Y.,
 Wakamatsu, A., Sugiyama, T., Nagai, K., Kojima, S., Otsuki, T. and
 Koga, H.
 Primers for synthesizing full length cDNA clones and their use
 Patent: EP 1396543-A 3469 10-MAR-2004;
 Research Association for Biotechnology (JP)

TITLE
 JOURNAL

FEATURES
source
CDS
Location/Qualifiers
1..2169
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
344..1543
/note="unnamed protein product"
/codon_start=1
/protein_id="CAF86540.1"
/db_xref="GI:45503260"
/translation="MAQGLVTPADVADIDPSQERMACTNSAQEDLYVDWLENYSNIVS
LDLESAYENKSLPTEKNIIHEIDASKNSDPSRSLGRNI CSEGLERPDRSGRYNQ
MLINYKRPATREBGPPTRHQRHKENSFECCGCAFRGQLSDHQKIHTEPEYCK
CECKKAFRWGNQLTQHKIHTEKEPECKDGKAFWSSIVIHRIHTEGPEYCK
DCGKAPRGDELTHORFHTGEKDYCKGCTFSRVYLIQHKRHSSEKPEYCKD
GKAFICGSSLIOHKRIHTEKPEYCKECCAKPTRVYLVLOHOKIHTEKPEYCKGK
ARWSSLYVKAHERIHTEKPEYCKTECGKAFNCGYHLTQHERIHTEKPEYCKGKAF
IVGSSLIV"

ORIGIN
Query Match 97.2%; Score 1958; DB 6; Length 2169;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 1961; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
47 CAGGCCAGCATCTTCAGAAAAAGCATCCCGAGAGAGAAAGCAATCGTTAAACATCTT 106
204 CAGGCCAGCATCTTCAGAAAAAGCATCCCGAGAGAGAAAGCAATCGTTAAACATCTT 263
QY 107 AGGTGAGCTTACCTCTCGGAATTTGTCTTCTTCAGTGAACCCCGAGAAACCTGATC 166
DB 264 AGGTGAGCTTACCTCTCGGAATTTGTCTTCTTCAGTGAACCCCGAGAAACCTGATC 323
QY 167 AGTTCTTCAGTTTAAACCAATGGCCAGGGTTTGTGACGTTCCGCCAGCTAGCCATAG 226
DB 324 AGTTCTTCAGTTTAAACCAATGGCCAGGGTTTGTGACGTTCCGCCAGCTAGCCATAG 383
QY 227 ACTTTTCTCAGAGAGAGTGGGCTGTCTGAACCTCTGCTCAGAGAGACCTTGACTGGAGC 286
DB 384 ACTTTTCTCAGAGAGAGTGGGCTGTCTGAACCTCTGCTCAGAGAGACCTTGACTGGAGC 443
QY 287 TGAATCTGAGAACTACGTAACCTTGTCTCACTGGAATTTGGAGTCAAGCATTAAGAAATA 346
DB 444 TGAATCTGAGAACTACGTAACCTTGTCTCACTGGAATTTGGAGTCAAGCATTAAGAAATA 503
QY 347 AGAGTTTACCTACAGAAAAAAACATTCATGAATTAAGGGCTTCCAAAAGGAATTCAGATA 406
DB 504 AGAGTTTACCTACAGAAAAAAACATTCATGAATTAAGGGCTTCCAAAAGGAATTCAGATA 563
QY 407 GAAAGATTAATCCCTTGGCCGTAACCTGATATGTGAAGGTACGCTTGAAAGACCAAGC 466
DB 564 GAAAGATTAATCCCTTGGCCGTAACCTGATATGTGAAGGTACGCTTGAAAGACCAAGC 623
QY 467 GCTCCAGAGGAGATATGTCAATCAAGTGAATCAATTAATGTCAAAAGGCCCTGCTACTA 526
DB 624 GCTCCAGAGGAGATATGTCAATCAAGTGAATCAATTAATGTCAAAAGGCCCTGCTACTA 683
QY 527 GAGAAGGCAACCCCTCTAGAAACAATCAGAGACATCATAGAGAAATTCCTTTGAATGTA 586
DB 684 GAGAAGGCAACCCCTCTAGAAACAATCAGAGACATCATAGAGAAATTCCTTTGAATGTA 743
QY 587 AGGACTGTGGAAAGGCTTTAGTCGTGGCTATCAACTTAAGTCAACATCAGAAAAATCCATA 646
DB 744 AGGACTGTGGAAAGGCTTTAGTCGTGGCTATCAACTTAAGTCAACATCAGAAAAATCCATA 803
QY 647 CTGGTGAAGAACTTATGAATGTAAAGATGTAAAGAGGCTTCCGTTGGGGCAATACG 706
DB 804 CTGGTGAAGAACTTATGAATGTAAAGATGTAAAGAGGCTTCCGTTGGGGCAATACG 863
QY 707 TTAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAAT 766
DB 864 TTAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAAT 923
QY 767 AGGCTTTTTCATGGGGCTCAAGGCTCGTTATTCATTAAGAGGATTCATATCGTGGAAAAAC 826

DB 924 AGGCTTTTTCATGGGGCTCAAGGCTCGTTATTCATTAAGAGGATTCATATCGTGGAAAAAC 983
QY 827 CCTATGAATGTAAAGACTGTGAAAGGCGCTTCCGGCGTGTATGAGCTCACTCAGACAC 886
DB 984 CCTATGAATGTAAAGACTGTGAAAGGCGCTTCCGGCGTGTATGAGCTCACTCAGACAC 1043
QY 887 AGAGATTCACACCTGGGAGAAAGACTAGAAATGCAAGAACTGTGGGAAAGACTTTTAGCC 946
DB 1044 AGAGATTCACACCTGGGAGAAAGACTAGAAATGCAAGAACTGTGGGAAAGACTTTTAGCC 1103
QY 947 GGTGTATTAATCTTATTCAGCAAGAGAAATTCATATGTGGGAGAAAGCTTACGAGTGA 1006
DB 1104 GGTGTATTAATCTTATTCAGCAAGAGAAATTCATATGTGGGAGAAAGCTTACGAGTGA 1163
QY 1007 AAGACTGTGGGAAAGGCTTTATTTGTGGTTCAAGGCTCATTCAGCATTAAGAAATTCACA 1066
DB 1164 AAGACTGTGGGAAAGGCTTTATTTGTGGTTCAAGGCTCATTCAGCATTAAGAAATTCACA 1223
QY 1067 CAGTGAAGAAACCTATGATGTCAAGATGTGGAGGCTTTTACTCGAGTCAATTAAC 1126
DB 1224 CAGTGAAGAAACCTATGATGTCAAGATGTGGAGGCTTTTACTCGAGTCAATTAAC 1283
QY 1127 TTAATCAGATCAAGAAATCCACACGGGTGAAGGCTTACGAAATGTAAAGAGTGGGA 1186
DB 1284 TTAATCAGATCAAGAAATCCACACGGGTGAAGGCTTACGAAATGTAAAGAGTGGGA 1343
QY 1187 AAGGCTTTGCTGGGGTTGAGGCTCGTTAAGACAGAGAGATCAATACGGGCGAGAAAC 1246
DB 1344 AAGGCTTTGCTGGGGTTGAGGCTCGTTAAGACAGAGAGATCAATACGGGCGAGAAAC 1403
QY 1247 CGTACAGTGAACAGAAATGTGGAAAGGCTTCAATTTGTGGCTATCACTCAGACAG 1306
DB 1404 CGTACAGTGAACAGAAATGTGGAAAGGCTTCAATTTGTGGCTATCACTCAGACAG 1463
QY 1307 AGAGAAATCCACACAGGCGAAACCCCGTATTAATGTAAAGAGTGTGGAAAGGCTTTCAATT 1366
DB 1464 AGAGAAATCCACACAGGCGAAACCCCGTATTAATGTAAAGAGTGTGGAAAGGCTTTCAATT 1523
QY 1367 ATGGATCGAGCCTCGTGAACATGAGAGAAATTCATCCGGGGTGAACCTTATGGTGTGA 1426
DB 1524 ATGGATCGAGCCTCGTGAACATGAGAGAAATTCATCCGGGGTGAACCTTATGGTGTGA 1583
QY 1427 CAGAAATGTGGAAAGACTTTAGTCAAGGCTATCAGCTTACACAACTCAGAAAAACGACA 1486
DB 1584 CAGAAATGTGGAAAGACTTTAGTCAAGGCTATCAGCTTACACAACTCAGAAAAACGACA 1643
QY 1487 GTGGGGCGAAATCTTACGAATGTAAAGAGTGGGGAAAGGCTTAACTCAACTTAACATC 1546
DB 1644 GTGGGGCGAAATCTTACGAATGTAAAGAGTGGGGAAAGGCTTAACTCAACTTAACATC 1703
QY 1547 TCCGAGAACATCAGAGAGATCCCAACAGTTGAAGAGCCTTTTGAAGCGAGTAGCCGCTC 1606
DB 1704 TCCGAGAACATCAGAGAGATCCCAACAGTTGAAGAGCCTTTTGAAGCGAGTAGCCGCTC 1763
QY 1607 GATATCTATGTTTGGCTTTCACAGTTTGTACCTGCAAGTCAACTGCACTTCAAAAATAT 1666
DB 1764 GATATCTATGTTTGGCTTTCACAGTTTGTACCTGCAAGTCAACTGCACTTCAAAAATAT 1823
QY 1667 TAAATGAAAAATTCAGAAAAATTAAGATTTTAAGTCTCAAATGTGTGCTTCTGAGTA 1726
DB 1824 TAAATGAAAAATTCAGAAAAATTAAGATTTTAAGTCTCAAATGTGTGCTTCTGAGTA 1883
QY 1727 GCGTGAATGAATCTCGGCTGTCGGGCTCAGCGGCGGGGAGTGAAGTCAATCCCTTGG 1786
DB 1884 GCGTGAATGAATCTCGGCTGTCGGGCTCAGCGGCGGGGAGTGAAGTCAATCCCTTGG 1943
QY 1787 TCCAGACATCCACGCTGTATAGCCACCACTCTGATGACTTAAGTACCGGTCTTGG 1846
DB 1944 TCCAGACATCCACGCTGTATAGCCACCACTCTGATGACTTAAGTACCGGTCTTGG 2003
QY 1847 TGAATCAATCAATTCACAGATCAAGTGTGCTGTGCGCAAGTATCTTCACTTTGCTT 1906

Db	2004	TGATGAGATCAACTATCCACAGATCCAGTGCCTGTGCCCAGACAGTCCCTCACTTGGTT	2063
Oy	1907	AACAGTGGCCCCAGAGACAGACAGTAGTGATGCTGGTGATTTCGATATGCGAAGAAG	1966
Db	2064	AACAGTGGCCCCAGAGACAGAGAGTGAAGTGTGGTGTGATTTCGATATGCGAAGAAG	2123
Oy	1967	CCACAAGTGTCTTCCTTTAATGAAAAGTGAAAGTTCCTCACTT	2012
Db	2124	CCACAAGTGTCTTCCTTTAATGAAAAGTGAAAGTTCCTCACTT	2169
RESULT 7			
AK075226		2169 bp	mRNA linear PRI 03-SEP-2002
DEFINITION		Human sapiens cDNA FLJ90745 f1s, clone PLACE1011407, moderately	
ACCESSION		AK075226	
VERSION		AK075226.1 GI:22761178	
KEYWORDS		oligo capping, f1s (full insert sequence).	
SOURCE		Homo sapiens (human)	
ORGANISM		Homo sapiens	
REFERENCE		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
AUTHORS		Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;	
		Hominidae; Homo.	
	1		
	Isogai,T., Ota,T., Nishikawa,T., Hayashi,K., Otsuki,T.,		
	Sugiyama,T., Suzuki,Y., Negai,K., Sugano,S., Ishii,S.,		
	Kawai-Hio,Y., Saito,K., Yamamoto,J., Wakamatsu,A., Nakamura,Y.,		
	Kojima,S., Negahari,K., Masuho,Y., Ono,T., Okano,K., Yoshikawa,Y.,		
	Aotsuka,S., Sasaki,N., Hattori,A., Okumura,K., Iwayanagi,T. and		
	Ninomiya,K.		
	NEDO human cDNA sequencing project		
	Unpublished		
	2 (bases 1 to 2169)		
TITLE		Isogai,T. and Otsuki,T.	
JOURNAL		Direct Submission	
AUTHORS		Submitted (25-MAR-2002) Takao Isogai, Helix Research Institute,	
		Genomics Laboratory, 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan	
		(E-mail:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)	
COMMENT		NEDO human cDNA sequencing project supported by Ministry of	
		Economy, Trade and Industry of Japan; cDNA full insert sequencing;	
		Research Association for Biotechnology; cDNA library construction;	
		Institute of Medical Science, University of Tokyo, Laboratory of	
		Genome Structure, Human Genome Center; cDNA 5'- & 3'-end one pass	
		sequencing and clone selection; Helix Research Institute (supported	
		by Japan Key Technology Center etc.).	
FEATURES		Location/Qualifiers	
source		1..2169	
		/organism="Homo sapiens"	
		/mol_type="mRNA"	
		/db_xref="taxon:9606"	
		/clone="PLACE1011407"	
		/tissue_type="Placenta"	
		/clone_lib="PLACE1"	
		/note="cloning vector: PME18SFL3"	
		344..1543	
		/note="unnamed protein product"	
		/codon_start=1	
		/protein_id="BAC11485.1"	
		/db_xref="GI:22761179"	
		/translation="MAOGLYTFADVAIDFSQEWACLNASQRLDLYMDVLMENYSLVS	
		LDLSAVENKSLPTEKNLHIFRASKRNSDRSKSLGNWICEGLRFPQSRGRVYVNOQ	
		MLINVYKPRNREGPTPRTHRRHKENSFEKDCGKA.PSRGYOLSHQKIHTEGKYRE	
		CKEKA.PRMGNOLTOHOKIHTEGKYPEDCCGCA.FPMGSLVYHKIHTEGKYRECK	
		DGCA.FRRGDELTOHOREHTEGKYRECKDCGKTSRYKLI.QHGRHSGEFVYCKCQ	
		GKATIGSSLI.QHGRHTEGKYRECKGCA.FTVVYVTOHOKIHTEGKYRECKGCA	
		AFRNGSSLYVKHRIHTEGKYRKYTCGCA.FNCGVYLTQHERIHTEGTYPKCKGCA	
		IYGSLSLV"	
CDS			
ORIGIN			
Query Match	97.2%	Score 1958;	DB 8; Length 2169;
Best Local Similarity	99.7%;	Prod. No. 0;	
Matches 1961;	Conservative	0;	Mismatches 5; Indels 0; Gaps 0;

QY	47	TAAGCCAGACATCTTTCAGAAAAAACATCCCGAGAGAGAAAGACAAATCGTTAAACATCTT	106
Db	204	CAGGCGACACATCTTTCAGAAAAAACATCCCGAGAGAGAAAGACAAATCGTTAAACATCTT	263
QY	107	AGGTCAGGCTCAGGCTCTCGGAAATTTGTCTCTTCAGTGGAAACCCCGAGAGACTGATC	166
Db	264	AGGTCAGGCTCAGGCTCTCGGAAATTTGTCTCTTCAGTGGAAACCCCGAGAGACTGATC	323
QY	167	AGTTCTTCAGTTCTPAAACAAATGCCCCAGGGGTTTGTGACGTTGCGGACGTAGCCATAG	226
Db	324	AGTTCTTCAGTTCTPAAACAAATGCCCCAGGGGTTTGTGACGTTGCGGACGTAGCCATAG	383
QY	227	ACTTTTCACAGAGAGAGGGGCGGTGTGAACTGTGCTCAGAGGGGACCTGTACTGGGACG	286
Db	384	ACTTTTCACAGAGAGAGGGGCGGTGTGAACTGTGCTCAGAGGGGACCTGTACTGGGACG	443
QY	287	TGATGCTGAGAACATCAGTAACCTTGTCTTCACCTGATTTTGGAGTCAGCATATGAAATA	346
Db	444	TGATGCTGAGAACATCAGTAACCTTGTCTTCACCTGATTTTGGAGTCAGCATATGAAATA	503
QY	347	AGAGTTTACCTACAGAAAAAAACATTCATGAAATTAAGGGCTTCCMAAAGAAATTCAGATA	406
Db	504	AGAGTTTACCTACAGAAAAAAACATTCATGAAATTAAGGGCTTCCMAAAGAAATTCAGATA	563
QY	407	GAAGAAGTAAATCCCTTGGCCGTAACTGATATGTGAAGGTAACCTTGAAGAACAAGC	466
Db	564	GAAGAAGTAAATCCCTTGGCCGTAACTGATATGTGAAGGTAACCTTGAAGAACAAGC	623
QY	467	GCTCCAGAGGAGGTATGTCAATCAGATGATCATCAATTAATGTCAAAAAGGCTTGCTA	526
Db	624	GCTCCAGAGGAGGTATGTCAATCAGATGATCATCAATTAATGTCAAAAAGGCTTGCTA	683
QY	527	GAGAAAGGACCCCTCCATGAAACATATAGAGACATCAATTAAGGAAATTCCTTTGAATGTA	586
Db	684	GAGAAAGGACCCCTCCATGAAACATATAGAGACATCAATTAAGGAAATTCCTTTGAATGTA	743
QY	587	AGGACTGTGGGAAAGGCTTTAGTGTGGGCTTCAACTAGTCAACATCAGAAAAATTCATA	646
Db	744	AGGACTGTGGGAAAGGCTTTAGTGTGGGCTTCAACTAGTCAACATCAGAAAAATTCATA	803
QY	647	CTGGTGAGAAACTTTATGATGTGAAGATGTAAAGAGGCTTCGTTGGGGCAATCAGC	706
Db	804	CTGGTGAGAAACTTTATGATGTGAAGATGTAAAGAGGCTTCGTTGGGGCAATCAGC	863
QY	707	TTACTCAACATCAAAAAAATTCATCTAGGGGGAAGCCCTACGAATGTAAAGACTGTGGGA	766
Db	864	TTACTCAACATCAAAAAAATTCATCTAGGGGGAAGCCCTACGAATGTAAAGACTGTGGGA	923
QY	767	AGGCTTTTCAGTGGGGCTCAAGCTCGTATTCATTAAGAGGATTCATACTGGTGA AAAAC	826
Db	924	AGGCTTTTCAGTGGGGCTCAAGCTCGTATTCATTAAGAGGATTCATACTGGTGA AAAAC	983
QY	827	CCCTATGAAATGTAAGAATCTGTGGAAGAGGCTTTGGCGCTGTGTGAAGACCTCACTAGACAC	886
Db	984	CCCTATGAAATGTAAGAATCTGTGTGGAAGAGGCTTTGGCGCTGTGTGAAGACCTCACTAGACAC	1043
QY	887	AGAGATTCACAACCTGTGGGAGAAAGACTACAGATGCAAGAACTGTGGGAGAGACTTTAGCC	946
Db	1044	AGAGATTCACAACCTGTGGGAGAAAGACTACAGATGCAAGAACTGTGGGAGAGACTTTAGCC	1103
QY	947	GTGTGTATAACTTATTCAGCAACAAGAAATTCATAGTGGGAGAGGCTTTACAGATGTGA	1006
Db	1104	GTGTGTATAACTTATTCAGCAACAAGAAATTCATAGTGGGAGAGGCTTTACAGATGTGA	1163
QY	1007	AAGACTGTGGGAAAGGCTTTATTTGTGTGTTCAAGCCCTCATTCAGCATPAAAGAAATTCACA	1066
Db	1164	AAGACTGTGGGAAAGGCTTTATTTGTGTGTTCAAGCCCTCATTCAGCATPAAAGAAATTCACA	1223
QY	1067	CAGGTGAGAAACCTTATGATGTCAAGATGTGGGAAAGGCTTTATCTCGATCAATTAATCC	1126
Db	1224	CAGGTGAGAAACCTTATGATGTCAAGATGTGGGAAAGGCTTTATCTCGATCAATTAATCC	1283

QY 1127 TTACTCAGCATGAGAAATCCACACCGGTGAGAGAGCCTCAAGATGTAAAGAGTGTGGG 1186
 DB 1284 TTACTCAGCATGAGAAATCCACACCGGTGAGAGAGCCTCAAGATGTAAAGAGTGTGGG 1343
 QY 1187 AAGCCTTTGGCTGGGGTTTGAAGCTGTGTAAGACAGAGAGATACATAGGGGAGAGAAC 1246
 DB 1344 AAGCCTTTGGCTGGGGTTTGAAGCTGTGTAAGACAGAGAGATACATAGGGGAGAGAAC 1403
 QY 1247 CGTACAGTGCACAGAAATGGGAGAGCCTTCAATTTGGCTGTACCTCACTCAGACAG 1306
 DB 1404 CGTACAGTGCACAGAAATGGGAGAGCCTTCAATTTGGCTGTACCTCAGACAG 1463
 QY 1307 AAGAAATCCACACAGCGAAACCCCGTATTAATGTAAAGAGTGTGGAGAGCTTTCAATT 1366
 DB 1464 AAGAAATCCACACAGCGAAACCCCGTATTAATGTAAAGAGTGTGGAGAGCTTTCAATT 1523
 QY 1367 ATGATCGAGCCTCGTGAACAATGAGAGATTTATACCGGGGGTGAACCTTATGGGTGA 1426
 DB 1524 ATGATCGAGCCTCGTGAACAATGAGAGATTTATACCGGGGGTGAACCTTATGGGTGA 1583
 QY 1427 CAGAAATGTGGAGAGAGCTTTAGTCAAGGCGCATCAGCTTACACAACATCAGAAACGACAC 1486
 DB 1584 CAGAAATGTGGAGAGAGCTTTAGTCAAGGCGCATCAGCTTACACAACATCAGAAACGACAC 1643
 QY 1487 GTGGGGCGAAATCTTAAGAAATGTAAAGAGTGTGGAGAGCGATGTAAACCTTAACCATC 1546
 DB 1644 GTGGGGCGAAATCTTAAGAAATGTAAAGAGTGTGGAGAGCGATGTAAACCTTAACCATC 1703
 QY 1547 TCCGAGAAATCAGAGAGATTCACACAGTTGAAGAGCCTTTGAAGCGAGTACCCGCTC 1606
 DB 1704 TCCGAGAAATCAGAGAGATTCACACAGTTGAAGAGCCTTTGAAGCGAGTACCCGCTC 1763
 QY 1607 GATATCATGCTTCGCTTCCACAGTTGTTACCTGAGTCAACCTGAGTTCAAAATAT 1666
 DB 1764 GATATCATGCTTCGCTTCCACAGTTGTTACCTGAGTCAACCTGAGTTCAAAATAT 1823
 QY 1667 TAAATGAAATTCAGAAATTAAGAAATTTAAAGTCTCAATGTGTGCTTCTGAGTA 1726
 DB 1824 TAAATGAAATTCAGAAATTAAGAAATTTAAAGTCTCAATGTGTGCTTCTGAGTA 1883
 QY 1727 GCGTATGAAATTCCTGCTGTGCTCGGCTCAGAGCGGCGGGAGTGTGATCATCCCTTGG 1786
 DB 1884 GCGTATGAAATTCCTGCTGTGCTCGGCTCAGAGCGGCGGGAGTGTGATCATCCCTTGG 1943
 QY 1787 TCCAGACATCCAGCGCTGTATACGCCACCCCTGCTAGTGAAGTGTAGTGGCTTTGG 1846
 DB 1944 TCCAGACATCCAGCGCTGTATACGCCACCCCTGCTAGTGAAGTGTAGTGGCTTTGG 2003
 QY 1847 TGATCAGATCAATATCCAGCATCAGATGCTGTGCGCAAGTGTCTCACTTGTCTT 1906
 DB 2004 TGATCAGATCAATATCCAGCATCAGATGCTGTGCGCAAGTGTCTCACTTGTCTT 2063
 QY 1907 AAGATGTGGCGCCAGAGAGAGAGTGTATGCTGTGATTTGGATTTGGCAAGAGAG 1966
 DB 2064 AAGATGTGGCGCCAGAGAGAGAGTGTATGCTGTGATTTGGATTTGGCAAGAGAG 2123
 QY 1967 CCAAGAGTGTCTCTTTTAAATGAAAGAGGTAAGTGTCAACTT 2012
 DB 2124 CCAAGAGTGTCTCTTTTAAATGAAAGAGGTAAGTGTCAACTT 2169

RESULT 8
 AX405913 2010 bp DNA linear PAT 14-JUN-2002
 LOCUS AX405913
 DEFINITION Sequence 328 from Patent WO0222860.
 ACCESSION AX405913
 VERSION AX405913.1 GI:21439265
 KEYWORDS
 ORGANISM
 Homo sapiens (human)
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Homidae; Homo.

REFERENCE
 1
 TANG, Y. T., LIU, C., ZHOU, P., ASUNDI, V., ZHANG, J., ZHAO, Q. A., REN, F.,
 XUE, A. J., YANG, Y., WEHRMAN, T. and DRIMANAC, R. T.
 Novel nucleic acids and polypeptides
 Patent: WO 0222660-A 328 21-MAR-2002;
 HYSD, INC. (US)
 FEATURES
 source
 1..2010
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"
 383..1774
 /note="unnamed protein product"
 /codon_start=1
 /protein_id="CAD35020.1"
 /db_xref="GI:21439266"
 /translation="MAQGLVTFADVAIDPSEBEMACINSABDLYMDVWLENYSNLS
 LDLESAYENKSLPTERKHITETASRNRNDRBSKSGRWICGTELEPDRSGRYNO
 MLINTVKRPEATGEPPTTHQHHKNSFECDCKGASRGYQLSOHQKIHGEPKYE
 CECKKAFRMGNQLTHQKIHGEPKYECDCKGKAFRMGSSLVHHRHIGEPKYE
 DCKAFRRDELTOHQRFTGKEDEYCDCKGCTFBRVYLIOHKIHGEPKYECDCK
 GKAFIIGSSLIQKRIHIGEPKYECDCKGKAFRVYLVLOHKIHGEPKYECDCK
 AFRMSSLVKHERIHGEPKYECDCKGKAFNGYHLOHERIHGEPKYECDCKGKAF
 IYSSLVKHERIHGEPKYECDCKGKAFNGYHLOHERIHGEPKYECDCKGKAF
 LNHLEKHORHNS"
 ORIGIN
 Query Match 87.7%; Score 1768; DB 6; Length 2010;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1768; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 47 CAGGCCAGCATCCTTCAGAAAAAGCATCCCGAGAGAGAAAGCAATCGTTAAACATCTT 106
 DB 243 CAGGCCAGCATCCTTCAGAAAAAGCATCCCGAGAGAGAAAGCAATCGTTAAACATCTT 302
 QY 107 AGGTCAAGCTTACCTCTCGGAATTTGTCTTCTCACTGTGAAACCCCGAGAGAGATC 166
 DB 303 AGGTCAAGCTTACCTCTCGGAATTTGTCTTCTCACTGTGAAACCCCGAGAGAGATC 362
 QY 167 AGTTCTTCAGTTCTTAAACATATGCGCCAGAGGTTTGTGACGTTGCGCGAGCTAGCATAG 226
 DB 363 AGTTCTTCAGTTCTTAAACATATGCGCCAGAGGTTTGTGACGTTGCGCGAGCTAGCATAG 422
 QY 227 ACTTTTCTCAGAGAGAGTGGGCTGTCTGAACCTCTGCTCAGAGAGACCTGTACTGGACG 286
 DB 423 ACTTTTCTCAGAGAGAGTGGGCTGTCTGAACCTCTGCTCAGAGAGACCTGTACTGGACG 482
 QY 287 TGATGCTGAGAACTACATTAATTTGTCTCACTGGAATTTGGATGTACGATATGAAGTA 346
 DB 483 TGATGCTGAGAACTACATTAATTTGTCTCACTGGAATTTGGATGTACGATATGAAGTA 542
 QY 347 AGAGTTTACCTACAGAAAAAACATTCATGAATTAAGGCTTCAAAAGGAATTCAGATA 406
 DB 543 AGAGTTTACCTACAGAAAAAACATTCATGAATTAAGGCTTCAAAAGGAATTCAGATA 602
 QY 407 GAAGAAATTAATCCTTGGCGCGTAACTGTATGTGAAGGTACGCTTGAAGACACAGC 466
 DB 603 GAAGAAATTAATCCTTGGCGCGTAACTGTATGTGAAGGTACGCTTGAAGACACAGC 662
 QY 467 GCTTCCAGAGGAGATATGTCAATCAGATGATCATTAATTTATGTCAAAAGGCTGTACTTA 526
 DB 663 GCTTCCAGAGGAGATATGTCAATCAGATGATCATTAATTTATGTCAAAAGGCTGTACTTA 722
 QY 527 GAGAGGACACCCCTCTAGAACATCAGAGACATATAGAGAAATTCCTTTGAATGTA 586
 DB 723 GAGAGGACACCCCTCTAGAACATCAGAGACATATAGAGAAATTCCTTTGAATGTA 782
 QY 587 AGGACTGTGGAGAGGCTTTAGTGTGTGCTATCACTTATGTCAACATCAGAAAAATCATA 646
 DB 783 AGGACTGTGGAGAGGCTTTAGTGTGTGCTATCACTTATGTCAACATCAGAAAAATCATA 842
 QY 647 CTGTGAGAAACCTTAATGAATGTAAAGATGTAAAGAGGCTTTCGTTGGGGCAATCAGC 706

```

Db      843 CTGTGAGAAACCTTATGATGTAAAGAAATGTAAGAGGCTTCCTGGGGCAATCAGC 902
Qy      707 TTAATCAACATCAAAAAATTTTACATCTGGGAGAAAGCCCTTACGAATTTAAAGCTGGGA 766
Db      903 TTACTCAACATCAAAAAATTTTACATCTGGGAGAAAGCCCTTACGAATTTAAAGCTGGGA 962
Qy      767 AGGCTTTTGGATGGGCTCAAGCCTTGTATTCTATTAAGAGATTCTATCTGGTAAAAAAC 826
Db      963 AGGCTTTTGGATGGGCTCAAGCCTTGTATTCTATTAAGAGATTCTATCTGGTAAAAAAC 1022
Qy      827 CCTATGATGTAAAGACTGTGAAAGGCTTTTCGGGTGTGATGAGCTCACTCAGCACC 886
Db      1023 CCTATGATGTAAAGACTGTGAAAGGCTTTTCGGGTGTGATGAGCTCACTCAGCACC 1082
Qy      887 AAGAGATTCACACTGGGGGAGAAAGACTAAGATTCGCAAAACTGTGGGAAGACTTTTAGCC 946
Db      1083 AAGAGATTCACACTGGGGGAGAAAGACTAAGATTCGCAAAACTGTGGGAAGACTTTTAGCC 1142
Qy      947 GTGTGTATTAACCTTATTCAGCACAAAGAAATTCATGTGGGGAGAAAGCCTTACGAGTGA 1006
Db      1143 GTGTGTATTAACCTTATTCAGCACAAAGAAATTCATGTGGGGAGAAAGCCTTACGAGTGA 1202
Qy      1007 AAGACTGTGGGAAGGCTTTTATTTGTGTCAAGCCTCATTCAGCATTAAGAAATTCACA 1066
Db      1203 AAGACTGTGGGAAGGCTTTTATTTGTGTCAAGCCTCATTCAGCATTAAGAAATTCACA 1262
Qy      1067 CAGGTGAGAAACCTTATGAAATGTCAAGAAATGTGGGAAGGCTTTTACTCGAGTCAATTACC 1126
Db      1263 CAGGTGAGAAACCTTATGAAATGTCAAGAAATGTGGGAAGGCTTTTACTCGAGTCAATTACC 1322
Qy      1127 TTAATCAGCATCAAGAAATTCACACCGGTGAGAAAGCTTCAGAAATGTAAGAGTGGGA 1186
Db      1323 TTAATCAGCATCAAGAAATTCACACCGGTGAGAAAGCTTCAGAAATGTAAGAGTGGGA 1382
Qy      1187 AGGCTTTTGTGCTGGGGTTCGAGCCTGTTAAGACAGAGATACATACGCGGAGAAAGC 1246
Db      1383 AGGCTTTTGTGCTGGGGTTCGAGCCTGTTAAGACAGAGATACATACGCGGAGAAAGC 1442
Qy      1247 CGTACAAAGTGCACAGAAATGTGGGAAGGCTTTCAATTGTGGCTATCACTCTCAGCAGC 1306
Db      1443 CGTACAAAGTGCACAGAAATGTGGGAAGGCTTTCAATTGTGGCTATCACTCTCAGCAGC 1502
Qy      1307 AAGAGATTCACACAGGCGGAAACCCGTATTAATGTAAAGAGTGGGAGAGGCTTTCAATT 1366
Db      1503 AAGAGATTCACACAGGCGGAAACCCGTATTAATGTAAAGAGTGGGAGAGGCTTTCAATT 1562
Qy      1367 ATGATTCGAGCCTCGTGAACAATGAGAGATTCATACCGGGGTGAACCTTATGGGTGA 1426
Db      1563 ATGATTCGAGCCTCGTGAACAATGAGAGATTCATACCGGGGTGAACCTTATGGGTGA 1622
Qy      1427 CAGAAATGTGGGAAGGCTTTTATGTCAGGCTCATCAGCTTACACAACTCAAGAAAGCACA 1486
Db      1623 CAGAAATGTGGGAAGGCTTTTATGTCAGGCTCATCAGCTTACACAACTCAAGAAAGCACA 1682
Qy      1487 GTGGGGCGAAATCTTACGAATGTAAAGAGTGGGGAAGGCAATGTAACAACCTTAACCAATC 1546
Db      1683 GTGGGGCGAAATCTTACGAATGTAAAGAGTGGGGAAGGCAATGTAACAACCTTAACCAATC 1742
Qy      1547 TCCGAGAAATCATGAGAGATTCACAACAAGTGAAGAGCCTTTTGAACGCAAGTACCGCTC 1606
Db      1743 TCCGAGAAATCATGAGAGATTCACAACAAGTGAAGAGCCTTTTGAACGCAAGTACCGCTC 1802
Qy      1607 GATATCTATGTGTTTCCGTTTCCACAGTTTGTATCTCTGCAATCACTGCAAGTTCAAAATAT 1666
Db      1803 GATATCTATGTGTTTCCGTTTCCACAGTTTGTATCTCTGCAATCACTGCAAGTTCAAAATAT 1862
Qy      1667 TAAATGAGAAATTCGAGAAATTAAGAAATTTTAAGTCTCAAAATGTGTGCTTCTGAGTA 1726
Db      1863 TAAATGAGAAATTCGAGAAATTAAGAAATTTTAAGTCTCAAAATGTGTGCTTCTGAGTA 1922
Qy      1727 GCGTATGAAATCTCTCGCTGTCCGGCTCCAGCCGGCCGGGAGTGTGAGTCAATCCCTTGG 1786
Db      1923 GCGTATGAAATCTCTCGCTGTCCGGCTCCAGCCGGCCGGGAGTGTGAGTCAATCCCTTGG 1982

```

```

Qy      1787 TCCAGACATCCAGCGGTGATACGCCAC 1814
Db      1983 TCCAGACATCCAGCGGTGATACGCCAC 2010

RESULT 9
AC011487/c
LOCUS      AC011487      138579 bp      DNA      linear      PRI 28-FEB-2001
DEFINITION Homo sapiens chromosome 19 clone CTB-16765, complete sequence.
ACCESSION AC011487
VERSION   AC011487.5 GI:13162511
KEYWORDS  HTG.
SOURCE    Homo sapiens (human)
ORGANISM  Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominae; Homo.
REFERENCE  1 (bases 1 to 138579)
            DOE Joint Genome Institute and Stanford Human Genome Center.
            Direct Submission
            Unpublished
            2 (bases 1 to 138579)
            DOE Joint Genome Institute.
            Direct Submission
            Submitted (07-OCT-1999) Production Sequencing Facility, DOE Joint
            Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
            3 (bases 1 to 138579)
            DOE Joint Genome Institute and Stanford Human Genome Center.
            Direct Submission
            Submitted (28-FEB-2001) DOE Joint Genome Institute, 2800 Mitchell
            Drive, Walnut Creek, CA 94598, USA
            On Feb 28, 2001 this sequence version replaced gi:7690126.
            Draft Sequence Produced by DOE Joint Genome Institute
            www.jgi.doe.gov
            www.sbgc.stanford.edu
            Finishing Completed at Stanford Human Genome Center
            www.sbgc.stanford.edu
            Quality: Phrap Quality >=40 100% of Sequence;
            Estimated Total Number of Errors is 0.
            STS Content:
            SHGC-5578 G14081
            WI-18576 G23919.

FEATURES             location/Qualifiers
     source           1..138579
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="19"
                     /clone="CTB-16765"

ORIGIN
Query Match      83.9%; Score 1690.8; DB 8; Length 138579;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1692; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy      322 GATTTGAGTCAAGCATTAAGAAATTAAGAGTTTAACTTACAGAAAAAAACATTATGAAATA 381
Db      67527 GATTTGAGTCAAGCATTAAGAAATTAAGAGTTTAACTTACAGAAAAAAACATTATGAAATA 67468
Qy      382 AGGCTTCCAAAAGGAATTCAGATAGAAAGTAATATCCCTTGGCCGTAACTGATATGT 441
Db      67467 AGGCTTCCAAAAGGAATTCAGATAGAAAGTAATATCCCTTGGCCGTAACTGATATGT 67408
Qy      442 GAAAGTACGCTTGAAGAAGCCACAGCGCTCCAGAGGAGGTATGTCAATCAGATGATCATC 501
Db      67407 GAAAGTACGCTTGAAGAAGCCACAGCGCTCCAGAGGAGGTATGTCAATCAGATGATCATC 67348
Qy      502 AATTATGTCAAAAGGCTGTACTAGAGAGCAACCTCTTGAAGACATAGAGACAT 561
Db      67347 AATTATGTCAAAAGGCTGTACTAGAGAGCAACCTCTTGAAGACATAGAGACAT 67288
Qy      562 CATAAGAGAAATTCCTTGAATGTAAAGACTGTGGGAAGGCTTTAGTGTGCTATCA 621

```

```

Db      67287 CATTAAGAGAAATTCCTTGAATGTAAAGACCTGTGGAGAGCCCTTGTGCTGTGCTATCA 67228
Qy      622  CTTAAGCAATCATGAAAAATCCATATCTGTGTAGAAACCTTATGATGTAAAGATGTAAAG 681
Db      67227 CTTAGTCAATCATGAAAAATCCATCTGTGTAGAAACCTTATGATGTAAAGATGTAAAG 67168
Qy      682  AAGGCTTCCGTTGGGGCAATCAGCTTACTCAACATCAAAAAATTCATCTGTGGAGAAAG 741
Db      67167 AAGGCTTCCGTTGGGGCAATCAGCTTACTCAACATCAAAAAATTCATCTGTGGAGAAAG 67108
Qy      742  CCTTACGAATGTAAACATGTGTGGAGAGCTTTTCATGTGGGGCTCAAGCCCTGTTATTCAT 801
Db      67107 CCTTACGAATGTAAACATGTGTGGAGAGCTTTTCATGTGGGGCTCAAGCCCTGTTATTCAT 67048
Qy      802  AAGAGGATTCATATCTGTGTGAAAAACCTTATGAATGTAAAGATGTGTGAAAAAGGCTTTCG 861
Db      67047 AAGAGGATTCATATCTGTGTGAAAAACCTTATGAATGTAAAGATGTGTGAAAAAGGCTTTCG 66988
Qy      862  CGTGTGTATGAGCTCATCTGACGACCAAGAGATTCACACTGTGGAGAAAGATCAAGATGC 921
Db      66987 CGTGTGTATGAGCTCATCTGACGACCAAGAGATTCACACTGTGGAGAAAGATCAAGATGC 66928
Qy      922  AAAAGACTGTGTGGAGAGACCTTACCGCTGTGTATTAACCTTATTCAGCACAAAGAAATTCAT 981
Db      66927 AAAAGACTGTGTGGAGAGACCTTACCGCTGTGTATTAACCTTATTCAGCACAAAGAAATTCAT 66868
Qy      982  AGTGGGAGAGAGCCTTACAGATGTAAAGATGTGTGGAGAGCCTTATTTGTGTTCAGC 1041
Db      66867 AGTGGGAGAGAGCCTTACAGATGTAAAGATGTGTGGAGAGCCTTATTTGTGTTCAGC 66808
Qy      1042 CTCATTCAGCATAAAGAAATTCACAGAGTGAGAAACCTTATGATGTCAAGATGTGG 1101
Db      66807 CTCATTCAGCATAAAGAAATTCACAGAGTGAGAAACCTTATGATGTCAAGATGTGG 66748
Qy      1102 AAGGCTTTCATCTGAGATCAATTCCTTACTCAGATCAGAAAGATTCACACCGGTGAGAA 1161
Db      66747 AAGGCTTTCATCTGAGATCAATTCCTTACTCAGATCAGAAAGATTCACACCGGTGAGAA 66688
Qy      1162 CCTCAGAAATGTAAAGAGTGTGGAGAGGCTTTCGCTGTGGGTTGAGCCCTGTTAAGCAC 1221
Db      66687 CCTCAGAAATGTAAAGAGTGTGGAGAGGCTTTCGCTGTGGGTTGAGCCCTGTTAAGCAC 66628
Qy      1222 GAGAGGATCATACGCGGCGAGAAAGCCGTACAGATGTGTGGAGAGGCTTTCAT 1281
Db      66627 GAGAGGATCATACGCGGCGAGAAAGCCGTACAGATGTGTGGAGAGGCTTTCAT 66568
Qy      1282 TGTGTCTATCACTCTCCTGACGAGAGAAATCCACACAGGCCAAACCCCGTATTAATGT 1341
Db      66567 TGTGTCTATCACTCTCCTGACGAGAGAAATCCACACAGGCCAAACCCCGTATTAATGT 66508
Qy      1342 AAGGAGTGTGTGGAGAGGCTTTCATTTATGATCGAGCCTCGTGAACATGAGAGATTCAT 1401
Db      66507 AAGGAGTGTGTGGAGAGGCTTTCATTTATGATCGAGCCTCGTGAACATGAGAGATTCAT 66448
Qy      1402 ACCGCGGTGAACCCCTATGTGGTGTACAGAAATGTGTGGAGAGGCTTTCATCGGCCATCAG 1461
Db      66447 ACCGCGGTGAACCCCTATGTGGTGTACAGAAATGTGTGGAGAGGCTTTCATCGGCCATCAG 66388
Qy      1462 CTTACCAAAATCATGAGAAAGGCAAGTGCGGGCGAAACCTTACGAATGTAAAGAGTGTGG 1521
Db      66387 CTTACCAAAATCATGAGAAAGGCAAGTGCGGGCGAAACCTTACGAATGTAAAGAGTGTGG 66328
Qy      1522 AAGGCAATGTAAACCACTTAACCATCTCCGAGAAATCATGAGAGATTCACAAAGTTGAAGA 1581
Db      66327 AAGGCAATGTAAACCACTTAACCATCTCCGAGAAATCATGAGAGATTCACAAAGTTGAAGA 66268
Qy      1582 GCGCTTTGAACGAGATGAGCCGCTCGTATCTATGATTTGCTTTCCACAGTTGTTCAT 1641
Db      66267 GCGCTTTGAACGAGATGAGCCGCTCGTATCTATGATTTGCTTTCCACAGTTGTTCAT 66208
Qy      1642 GCGATCAATCGCAGTTCAAAAAATTTAAATGTAAAAATTTCCAGAAATTAAGATTTTAAGT 1701
Db      66207 GCGATCAATCGCAGTTCAAAAAATTTAAATGTAAAAATTTCCAGAAATTAAGATTTTAAGT 66148

```

```

Qy      1702 CTCAAATGTGTGTGCCCTTCTGAGTAGCGTATGAATCTCTGCTGTCCGCTCCAGCCG 1761
Db      66147 CTCAAATGTGTGTGCCCTTCTGAGTAGCGTATGAATCTCTGCTGTCCGCTCCAGCCG 66088
Qy      1762 GCCGGGAGATGTAGATCATCCCTTGGTCCAGGACATCCAGCGCTGTATAGCCCAACCCCT 1821
Db      66087 GCCGGGAGATGTAGATCATCCCTTGGTCCAGGACATCCAGCGCTGTATAGCCCAACCCCT 66028
Qy      1822 GCTAGTAGCTTAGTATGAGCCGCTTGTGTGATCAGATCAATATCCAGCATCAAGTGCCTG 1881
Db      66027 GCTAGTAGCTTAGTATGAGCCGCTTGTGTGATCAGATCAATATCCAGCATCAAGTGCCTG 65968
Qy      1882 TGCCCAAGTAGTCTTCACCTTGTCTTAACAGTGGCCCCAGAGAGCAGAGTATGATGCTG 1941
Db      65967 TGCCCAAGTAGTCTTCACCTTGTCTTAACAGTGGCCCCAGAGAGCAGAGTATGATGCTG 65908
Qy      1942 GTGATTCGATATATGCCAAAGAGAGCCAAAGATGCTTCTTTAAATGAAAAAGTGAAA 2001
Db      65907 GTGATTCGATATATGCCAAAGAGAGCCAAAGATGCTTCTTTAAATGAAAAAGTGAAA 65848
Qy      2002 GTTCTCAACTTAAA 2015
Db      65847 GTTCTCAACTTAAA 65834

RESULT 10
AF251515      1904 bp      mRNA      linear      PRI 20-MAR-2001
LOCUS      Homo sapiens C2H2-like zinc finger protein (ZNF463) mRNA, complete
DEFINITION
ACCESSION      AF251515
VERSION      AF251515.2 GI:10048469
KEYWORDS
SOURCE
ORGANISM      Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominoidea; Homo.
1 (bases 1 to 1904)
Wu,H., Zhang,S., Qiu,W., Zhang,G., Xia,O., Xiao,C., Huang,X.,
Huang,M., Agen,F., Fan,T., Yang,J. and Milunsky,A.
Isolation, characterization, and mapping of a novel human KRB zinc
finger protein encoding gene ZNF463
Biochim. Biophys. Acta 1518 (1-2), 190-193 (2001)
11267678
2 (bases 1 to 1904)
Wu,H., Qiu,W., Zhang,S., Zhang,G. and Xiao,C.
Cloning of azoospermia-related genes
Unpublished
3 (bases 1 to 1904)
Wu,H., Qiu,W., Zhang,S., Zhang,G. and Xiao,C.
Direct Submission
Submitted (03-APR-2000) Medical Genetics, West China University of
Medical Sciences, Reminmananlu 3duan 17, Chengdu, Sichuan 610041,
China
4 (bases 1 to 1904)
Wu,H., Qiu,W., Zhang,S., Zhang,G. and Xiao,C.
Direct Submission
Submitted (11-SEP-2000) Medical Genetics, West China University of
Medical Sciences, Reminmananlu 3duan 17, Chengdu, Sichuan 610041,
China
REMARK
COMMENT      Sequence update by submitter
On Sep 11, 2000 this sequence version replaced gi:7804915.
FEATURES
Source
1..1904
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="19"
/map="19q13; between STS3417 and D19S425-D19S418"
/tissue_type="testis"
1..1904
gene

```


CDS

/gene="ZNF463"
 377..1768
 /gene="ZNF463"
 /note="similar to Kruppel-type zinc finger genes;
 potentially associated with spermatogenesis"
 /codon_start=1
 /product="C2H2-like zinc finger protein"
 /protein_id="AA070179.2"
 /db_xref="GI:10048470"
 /translation="MAQGLVTFADVADVAIPDSOEWMACINSAQRLIYDWLVLENYSNLVS
 LDEBAYENKSLPTEKNLHEIRASRNSDRSKSLGRWICEGLTERPQSRGRVYNQ
 MIINVKRPARREGETPRTHORHNKENSEFECCKAKSRGYLSLOHOKIHTEKPYE
 CECKKAFRMGNOLTOHOKIHTEKPYECCDCKAFRSGSLVIHRIHTGKPYECK
 DCKAFRRGDELTOHOKIHTEKPYECCDCKAFRSGSLVIHRIHTGKPYECK
 GKAPICGSSLILOHOKIHTEKPYECCDCKAFRSGSLVIHRIHTGKPYECK
 AKRWSLSLVKHERIHTGKPYKTECGAFNGYHLTQHERIHTGKPYECK
 IYGSVLVHERIHTGKPYKTECGAFNGYHLTQHERIHTGKPYECK
 LNLHREHORINS"

ORIGIN

Query Match 82.4%; Score 1659.4; DB 8; Length 1904;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 1660; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

47 CAGGCCAGCATCCTTCAGAAAAAGCATCCCGAGGAGGAAGCAATGTTAAACATCTT 106
 |||||
 237 CAGGCCAGCATCCTTCAGAAAAAGCATCCCGAGGAGGAAGCAATGTTAAACATCTT 296
 |||||
 107 AGGTCAGCTTAGCCTCTCGAATTTGTTCTTTCAGTGGAAACCCCGAAGACTGATC 166
 |||||
 297 AGGTCAGCTTAGCCTCTCGAATTTGTTCTTTCAGTGGAAACCCCGAAGACTGATC 356
 |||||
 167 AGTTCTTCACTTAAACAATAGCCCAAGGTTTGGTGAAGTTCCGCCAGCTGAACCAT 226
 |||||
 357 AGTTCTTCACTTAAACAATAGCCCAAGGTTTGGTGAAGTTCCGCCAGCTGAACCAT 416
 |||||
 227 ACTTTCTCAGGAGAGTGGGCTGTCTGAACCTCTGCTCAGAGGAGACCTGTCTGGGACG 286
 |||||
 417 ACTTTCTCAGGAGAGTGGGCTGTCTGAACCTCTGCTCAGAGGAGACCTGTCTGGGACG 476
 |||||
 287 TGATCTCGAGAACTACAGTAAGTGTCTCACTGGAATTTGGAGTCAGCATATGAATA 346
 |||||
 477 TGATCTCGAGAACTACAGTAAGTGTCTCACTGGAATTTGGAGTCAGCATATGAATA 536
 |||||
 347 AGAGTTTACTTACAGAAAAAAACATTCATGAATAAGGCTTCCAAAAGAAATTCAGATA 406
 |||||
 537 AGAGTTTACTTACAGAAAAAAACATTCATGAATAAGGCTTCCAAAAGAAATTCAGATA 596
 |||||
 407 GAAGAAGTAATCCCTTGGCCGTAAGTGAATGTGAAGGTACGCTTGAAAAGACCAACAGC 466
 |||||
 597 GAAGAAGTAATCCCTTGGCCGTAAGTGAATGTGAAGGTACGCTTGAAAAGACCAACAGC 656
 |||||
 467 GCTCCAGAGGAGATGATCATCATCATCATCATCATCATCATCATCATCATCATCATCAT 526
 |||||
 657 GCTCCAGAGGAGATGATCATCATCATCATCATCATCATCATCATCATCATCATCATCAT 716
 |||||
 527 GAGAAAGCAACCCCTCTTGAACAACATCAAGACATCATTAAGAGAAATTCCTTTGAATGA 586
 |||||
 717 GAGAAAGCAACCCCTCTTGAACAACATCAAGACATCATTAAGAGAAATTCCTTTGAATGA 776
 |||||
 587 AGGACTGTGGGAAAGGCTTTAGTGTGGCTATCACTTAAGTGAATTCAGAAAAATCCATA 646
 |||||
 777 AGGACTGTGGGAAAGGCTTTAGTGTGGCTATCACTTAAGTGAATTCAGAAAAATCCATA 836
 |||||
 647 CTGGTGAAGAACTTATGATGTAAGAAATGTGAAGAGCCCTTCGTTGGGGCAATCAGC 706
 |||||
 837 CTGGTGAAGAACTTATGATGTAAGAAATGTGAAGAGCCCTTCGTTGGGGCAATCAGC 896
 |||||
 707 TTACTCAACATCAAAAAATTCATCTGGGGGAGAAAGCCCTTCGAAATGTAAAGCTGTGGGA 766
 |||||
 897 TTACTCAACATCAAAAAATTCATCTGGGGGAGAAAGCCCTTCGAAATGTAAAGCTGTGGGA 956
 |||||
 767 AGGCTTTTCGATGGGGCTCAAGCCCTGTTATTCATTAAGAGATTCATCTGGTGAAGAAAC 826
 |||||

Db 957 AGGCTTTTCGATGGGGCTCAAGCCCTGTTATTCATTAAGAGATTCATCTGGTGAAGAAAC 1016
 |||||
 827 CCTATGAATGTAAAGACTGTGGAAAGCCCTTTGGCGGTGTGATAGCTCACTCAGACAC 886
 |||||
 1017 CCTATGAATGTAAAGACTGTGGAAAGCCCTTTGGCGGTGTGATAGCTCACTCAGACAC 1076
 |||||
 887 AGAGATTCACACTGGGGAGAAAGCTAAGATGAAGAACTGTGGAAAGCTTTAGCC 946
 |||||
 1077 AGAGATTCACACTGGGGAGAAAGCTAAGATGAAGAACTGTGGAAAGCTTTAGCC 1136
 |||||
 947 GTGTGTATAAATTATTCAGACAAGAAATTCATAGTGGGAGAAAGCTTTACAGATGA 1006
 |||||
 1137 GTGTGTATAAATTATTCAGACAAGAAATTCATAGTGGGAGAAAGCTTTACAGATGA 1196
 |||||
 1007 AAGACTGTGGGAAAGGCTTTATTTGTGTGTTCAAGCTCATTCAGATTAAGAAATTCACA 1066
 |||||
 1197 AAGACTGTGGGAAAGGCTTTATTTGTGTGTTCAAGCTCATTCAGATTAAGAAATTCACA 1256
 |||||
 1067 CAGGTGAAGAAACCTTATGAATGTCAAGAAATGTGGGAAAGGCTTTTCTGAGTCAATTAC 1126
 |||||
 1257 CAGGTGAAGAAACCTTATGAATGTCAAGAAATGTGGGAAAGGCTTTTCTGAGTCAATTAC 1316
 |||||
 1127 TTACTCAGCATCAGAAAGATCCACACCGGTGAGAAAGCTTCAGAAATGTAAAGAGTGGGA 1186
 |||||
 1317 TTACTCAGCATCAGAAAGATCCACACCGGTGAGAAAGCTTCAGAAATGTAAAGAGTGGGA 1376
 |||||
 1187 AGGCTTTTCGCTGGGGTTTCAAGCTCGTTAAGACAAGAGATACATACGGGCGAAGAC 1246
 |||||
 1377 AGGCTTTTCGCTGGGGTTTCAAGCTCGTTAAGACAAGAGATACATACGGGCGAAGAC 1436
 |||||
 1247 CGTACAAATGCAAGAAATGTGGGAAAGGCTTTCAATTGTGGCTATCACTCCTCAGACAG 1306
 |||||
 1437 CGTACAAATGCAAGAAATGTGGGAAAGGCTTTCAATTGTGGCTATCACTCCTCAGACAG 1496
 |||||
 1307 AGAGAAATCCACACAGGCGAAACCCCGTAAATGTAAAGAGTGTGGGAAAGGCTTTCAATT 1366
 |||||
 1497 AGAGAAATCCACACAGGCGAAACCCCGTAAATGTAAAGAGTGTGGGAAAGGCTTTCAATT 1556
 |||||
 1367 ATGATTCAGAGCTCGTGAACAATGAGAAATTCATACCGGGGTGAACCTTATGGGTGA 1426
 |||||
 1557 ATGATTCAGAGCTCGTGAACAATGAGAAATTCATACCGGGGTGAACCTTATGGGTGA 1616
 |||||
 1427 CAGAAATGTGGGAAAGGCTTTTACTCAGCGCCATCAGCTTACACATCAAGAAAGCAGACA 1486
 |||||
 1617 CAGAAATGTGGGAAAGGCTTTTACTCAGCGCCATCAGCTTACACATCAAGAAAGCAGACA 1676
 |||||
 1487 GTGGGGCGAAATCCCTAAGAAATGTAAAGAGTGTGGGAAAGGCAATGTAAACCACTTAAACATC 1546
 |||||
 1677 GTGGGGCGAAATCCCTAAGAAATGTAAAGAGTGTGGGAAAGGCAATGTAAACCACTTAAACATC 1736
 |||||
 1547 TCCGAGAAACATCAGAGATTCACACACAGTTGAAGAGCTTTTGAACGAGTAGCCGCTC 1606
 |||||
 1737 TCCGAGAAACATCAGAGATTCACACACAGTTGAAGAGCTTTTGAACGAGTAGCCGCTC 1796
 |||||
 1607 GTATCTATGTTTGGCTTTTCCACAGTTTGTATCTGCACTCACTCAGCTTCAAAAATAT 1666
 |||||
 1797 GTATCTATGTTTGGCTTTTCCACAGTTTGTATCTGCACTCACTCAGCTTCAAAAATAT 1856
 |||||
 1667 TAAATGGAATAATTCAGAAATTAAGAAATTTTAAGTCTCAAA 1707
 |||||
 1857 TAAATGGAATAATTCAGAAATTAAGAAATTTTAAGTCTCAAA 1897
 |||||

RESULT 11

AK128736 2838 bp mRNA linear PRI 30-JAN-2004
 LOCUS AK128736
 DEFINITION Homo sapiens cDNA FLJ16775 fis, clone BRCAN2020100, highly similar
 to Homo sapiens zinc finger protein 463 (ZNF463).
 ACCESSION AK128736
 VERSION AK128736.1 GI:34536261
 KEYWORDS oligo cloning; fis (full insert sequence).
 SOURCE Homo sapiens (human)

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homidae; Homo.

REFERENCE 1
AUTHORS Ota, T., Suzuki, Y., Nishikawa, T., Otsuki, T., Sugiyama, T., Irie, R.,
Kamakura, A., Hayashi, K., Sato, H., Nagai, K., Kimura, T., Irie, R.,
Sekine, M., Oobayashi, M., Nishi, T., Shibahara, T., Tanaka, T.,
Ishii, S., Yamamoto, J., Saito, K., Kawai, Y., Isono, Y., Nakamura, Y.,
Nagahari, K., Murakami, K., Yasuda, T., Iwayanagi, T., Wagatsuma, M.,
Shiratani, A., Sudo, H., Hosoiri, T., Kaku, Y., Kodaira, H., Kondo, H.,
Sugawara, M., Takahashi, M., Kanda, K., Yokoi, T., Furuya, T.,
Kikkawa, E., Omura, Y., Abe, K., Kamihara, K., Katsuta, N., Sato, K.,
Tanikawa, M., Yamazaki, M., Ninomiya, K., Ishibashi, T., Yamashita, H.,
Murakawa, K., Fujimori, K., Tanai, H., Kimata, M., Watanabe, H.,
Hiroda, M., Chiba, Y., Ishida, S., Ono, Y., Takiguchi, S., Watanabe, S.,
Yoshida, M., Horiuchi, T., Kusanagi, J., Kanehori, K., Takahashi-Fujii, A.,
Hara, H., Tanase, T., Nomura, Y., Togiya, S., Komai, F., Hara, R.,
Takeuchi, K., Arita, M., Imose, N., Musashino, K., Yuki, H., Oshima, A.,
Sasaki, N., Aotsuka, S., Yoshioka, Y., Matsunawa, H., Ichihara, T.,
Shichata, N., Sano, S., Moriya, S., Momiyama, H., Satoh, N., Takami, S.,
Terashima, Y., Suzuki, O., Nakagawa, S., Senoh, A., Mizoguchi, H.,
Goto, Y., Shimizu, F., Wakebe, H., Hishigaki, H., Watanabe, T.,
Sugiyama, A., Takemoto, M., Kawakami, B., Yamazaki, M., Watanabe, K.,
Kumagai, A., Itakura, S., Fukuzumi, Y., Fujimori, Y., Komiyama, M.,
Tashiro, H., Tanigami, A., Fujiwara, T., Ono, T., Yamada, K., Fujii, Y.,
Ozaki, K., Hiro, M., Ohmori, Y., Kawabata, A., Hikiji, T., Kobatake, N.,
Nagasaki, H., Ikema, Y., Okamoto, S., Ohtani, R., Kawakami, T.,
Noguchi, S., Itoh, T., Shigeta, K., Senba, T., Matsumura, K.,
Nakajima, Y., Mizuno, T., Morinaga, M., Sasaki, M., Togashi, T.,
Oyama, M., Hata, H., Watanabe, M., Komatsu, T., Mizushima-Sugano, J.,
Sato, T., Shirai, Y., Takahashi, Y., Nakagawa, K., Okumura, K.,
Nagase, T., Nomura, N., Kikuchi, H., Masuho, Y., Yamashita, R.,
Nakai, K., Yada, T., Nakamura, Y., Ohara, O., Isogai, T. and Sugano, S.
Complete sequencing and characterization of 21,243 full-length
human cDNAs
Nat. Genet. 36 (1), 40-45 (2004)
14702039

TITLE 2
JOURNAL Tanigami, A., Fujiwara, T., Shibahara, T., Goto, Y., Hiro, M.,
PUBMED Shintzu, F., Wakebe, H., Ono, T., Hishigaki, H., Watanabe, T., Ozaki, K.,
AUTHORS Sugiyama, T., Irie, R., Otsuki, T., Sato, H., Kamakura, A., Ishii, S.,
Yamamoto, J., Isono, Y., Kawai, H., Saito, K., Nishikawa, T.,
Kimura, K., Yamashita, H., Matsuo, K., Nakamura, Y., Sekine, M.,
Kikuchi, H., Kanda, K., Wagatsuma, M., Murakawa, K., Kanehori, K.,
Takahashi-Fujii, A., Oshima, A., Sugiyama, A., Kawakami, B., Suzuki, Y.,
Sugano, S., Nagahari, K., Masuho, Y., Nagai, K. and Isogai, T.
NEDO human cDNA sequencing project
Unpublished

TITLE 3 (bases 1 to 2838)
JOURNAL Isogai, T. and Yamamoto, J.
REFERENCE Direct Submission
AUTHORS Submitted (15-JUL-2003) Takao Isogai, FLJ Project (HRI Team); 2-6-7
JOURNAL Kazusa-Kamatari, Kisarazu, Chiba 292-0818, Japan
(E-mail: genomic@nri.co.jp, Tel: 81-438-52-3975, Fax: 81-438-52-3986)
COMMENT NEDO human cDNA sequencing project supported by Ministry of
Economy, Trade and Industry of Japan; cDNA full insert sequencing:
Research Association for Biotechnology (RAB); cDNA library
construction: Helix Research Institute (HRI) (supported by Japan
key Technology Center etc.); 5- & 3-end one pass sequencing: RAB,
HRI, and Biotechnology Center, National Institute of Technology and
Evaluation; clone selection for full insert sequencing: HRI and
RAB; annotation: Reverse Proteomics Research Institute, HRI and
RAB.

FEATURES 1. location/Qualifiers
source 1..2838
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="BRCAN2020100"
/cissue_type="candidate nucleus"
/clone_id="BRCAN2"
/note="cloning vector: pME18SFL3"

ORIGIN
Query Match 81.7%; Score 1645.4; DB 8; Length 2838;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1657; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

47 CAGGCCAGATCTCTTCAGAAAAAGCATCCCGAGAGGAAAGCGAATCTTAAATCTT 106
1180 CAGGCCAGATCTCTTCAGAAAAAGCATCCCGAGAGGAAAGCGAATCTTAAATCTT 1239
107 AAGTCAGCTCTAGCCCTCTGGAATTTGCTTCTTCAGTGAAGAACCCGGAAGACTGATC 166
1240 AAGTCAGCTCTAGCCCTCTGGAATTTGCTTCTTCAGTGAAGAACCCGGAAGACTGATC 1299
167 AGTTCCTCAGTCTTAAACAAATGAGCCAGAGGTTGGTGAAGTTCGCGAGCTGACCATG 226
1300 AGTTCCTCAGTCTTAAACAAATGAGCCAGAGGTTGGTGAAGTTCGCGAGCTGACCATG 1359
227 ACTTTTCTCAGAGAGAGTGGGCTGTCTGCACTCTGCTCAGAGGACCTGTACTGGAGAC 286
1360 ACTTTTCTCAGAGAGAGTGGGCTGTCTGCACTCTGCTCAGAGGACCTGTACTGGAGAC 1419
287 TGATGCTGGAAGACTAGTAACCTTGTCTCAGTGAATTGGAGTCAGCATATGAATAA 346
1420 TGATGCTGGAAGACTAGTAACCTTGTCTCAGTGAATTGGAGTCAGCATATGAATAA 1479
347 AGAGTTTACCTACAG-AAAAAACAATTCATGAATTAAGGGCTTCCAAAAGAAATTCAGAT 405
1480 AGAGTTTACCTACAGAAAAAACAATTCATGAATTAAGGGCTTCCAAAAGAAATTCAGAT 1539
406 AGAAGAGTAAATCCCTTGGCCGCTAATCTGATATGTGAAGGTACGTTGAAGACCAACG 465
1540 AGAAGAGTAAATCCCTTGGCCGCTAATCTGATATGTGAAGGTACGTTGAAGACCAACG 1599
466 CGCTCCAGAGGAGGATGATGATCATGATGATCATCAATTAATGTCAAAGGCTGCTACT 525
1600 CGCTCCAGAGGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1659
526 AGAAGAGGACCCCTCTAGAAACATCATGAGACATCATTAAGAGAAATTCCTTGAATGT 585
1660 AGAAGAGGACCCCTCTAGAAACATCATGAGACATCATTAAGAGAAATTCCTTGAATGT 1719
586 AAGACTGTGGGAGAGGCTTTTGTGTGCTGATCACTTAATGTCACATCAAGAAATTCAT 645
1720 AAGACTGTGGGAGAGGCTTTTGTGTGCTGATCACTTAATGTCACATCAAGAAATTCAT 1779
646 ACTGTGAGAAACCTTAATGATTAAGATTAAGAAATTAAGAAAGGCTTCCGTTGGGCAATCAG 705
1780 ACTGTGAGAAACCTTAATGATTAAGATTAAGAAATTAAGAAAGGCTTCCGTTGGGCAATCAG 1839
706 CTTAATCAACATCAAAAAATTCATATCTGGGAGAAAGCCCTAAGATTAAGATCTGGG 765
1840 CTTAATCAACATCAAAAAATTCATATCTGGGAGAAAGCCCTAAGATTAAGATCTGGG 1899
766 AAGGCTTTTCATGGGGCTCAAGCTCTGTTATTCATTAAGAGATTAATCTGTGAAAAA 825
1900 AAGGCTTTTCATGGGGCTCAAGCTCTGTTATTCATTAAGAGATTAATCTGTGAAAAA 1959
826 CCTATGATTAAGAAACATGAGGAAAGGCTTCCGCGGTGATGATGATGATGATGATGATGAT 885
1960 CCTATGATTAAGAAACATGAGGAAAGGCTTCCGCGGTGATGATGATGATGATGATGATGAT 2019
886 CAGAGATTCACACTGGGAGAAAGACTAGATGCAAGAGACTGTGGAGAAAGCTTTAGC 945
2020 CAGAGATTCACACTGGGAGAAAGACTAGATGCAAGAGACTGTGGAGAAAGCTTTAGC 2079
946 CGTGTATTAATTAATTCAGCAAGAGAAATTCATATGAGGAGAAAGCTTTAGCACTGT 1005
2080 CGTGTATTAATTAATTCAGCAAGAGAAATTCATATGAGGAGAAAGCTTTAGCACTGT 2139
1006 AAGACTGTGGAGAGGCTTTTATTTGTGTTCAAGGCTCATTCAGATTAAGAAATTCAC 1065
2140 AAGACTGTGGAGAGGCTTTTATTTGTGTTCAAGGCTCATTCAGATTAAGAAATTCAC 2199

QY 1066 ACAGGTGAGAAACCTATGAATGTCAAGATGTGGAGAGCCCTTACTCGACTCAATTAC 1125
 Db 2200 AAGGTGAGAAACCTATGAATGTCAAGATGTGGAGAGCCCTTACTCGACTCAATTAC 2259
 QY 1126 CTTACTCAGCATCAGAAAGTCCACACCGGTGAGAAAGCTCAGAAATGTAGAGTGTGG 1185
 Db 2260 CTTACTCAGCATCAGAAAGTCCACACCGGTGAGAAAGCTCAGAAATGTAGAGTGTGG 2319
 QY 1186 AAGGCTTTCCGTGGGTTTCAGAGCTCTGTTAAGCAGAGAGATACATACGGGCGAAG 1245
 Db 2320 AAGGCTTTCCGTGGGTTTCAGAGCTCTGTTAAGCAGAGAGATACATACGGGCGAAG 2379
 QY 1246 CCGTCAAGTGCAGAAATGTGGAGAGCCCTTCAATTGTGCTATCACTCACTGACAC 1305
 Db 2380 CCGTCAAGTGCAGAAATGTGGAGAGCCCTTCAATTGTGCTATCACTCACTGACAC 2439
 QY 1306 GAGAGATCCACACAGGCGAAACCCGTAATAATGTAGAGATGTGGAGAGCTTTCATT 1365
 Db 2440 GAGAGATCCACACAGGCGAAACCCGTAATAATGTAGAGATGTGGAGAGCTTTCATT 2499
 QY 1366 TATGATCGAGCTCGTGAACATGAGAGATTCATACCGGGGTGAAACCTATGAGTGT 1425
 Db 2500 TATGATCGAGCTCGTGAACATGAGAGATTCATACCGGGGTGAAACCTATGAGTGT 2559
 QY 1426 ACAGAAATGTGGAGAGCTTGTACAGCGCCATCAGCTTACACACATCAGAAACGAC 1485
 Db 2560 ACAGAAATGTGGAGAGCTTGTGTACAGCGCCATCAGCTTACACACATCAGAAACGAC 2619
 QY 1486 AGTGGGCGAAATCTTACGAATGTAGAGATGTGGAGAGCGCATGTACCACTTAACCAT 1545
 Db 2620 AGTGGGCGAAATCTTACGAATGTAGAGATGTGGAGAGCGCATGTACCACTTAACCAT 2679
 QY 1546 CTCCGAGAAATCAGAGATTCACAAAGTTGAGAGCTTTGAGACGAGTACCGCT 1605
 Db 2680 CTCCGAGAAATCAGAGATTCACAAAGTTGAGAGCTTTGAGACGAGTACCGCT 2739
 QY 1606 CGTATCTATGTGTTGGCTTTCACAGTTTGTACCTGCACTGCACTGCACTGCACTG 1665
 Db 2740 CGTATCTATGTGTTGGCTTTCACAGTTTGTACCTGCACTGCACTGCACTGCACTG 2799
 QY 1666 TTAATGAGAAATTCAGAAATTAAGAAATTTTAAGTCTC 1704
 Db 2800 TTAATGAGAAATTCAGAAATTAAGAAATTTTAAGTCTC 2838

RESULT 12
 CS118206 2196 bp DNA linear PAT 08-JUL-2005
 LOCUS CS118206
 DEFINITION Sequence 253 from Patent WO2005054508.
 ACCESSION CS118206 GI:70666152
 VERSION CS118206.1
 KEYWORDS
 SOURCE
 ORGANISM
 OTHER SEQUENCES; artificial sequences.
 REFERENCE
 AUTHORS 1
 TITLE Bertucci, F., Houlgate, R., Birnbaum, D. and Debono, S.
 JOURNAL Gene expression profiling of colon cancer by dna microarrays and correlation with survival and histoclinical parameters
 Patent: WO 2005054508-A 253 16-JUN-2005;
 Ipsogen (FR); Institut Paoli-Calmettes, Igc (FR); Institut National de la Sante et de la Recherche Medicale (INSERM) (FR)
 FEATURES
 source 1..2196
 /organism="synthetic construct"
 /mol_type="unassigned DNA"
 /db_xref="taxon:32630"
 /note="Description of Artificial sequences:primer"
 misc_feature 1..2196
 /note="zinc finger protein 463 (ZNF463) gene."

ORIGIN
 misc_feature 1..2196
 /note="zinc finger protein 463 (ZNF463) gene."

Query Match 80.4%; Score 1619.4; DB 6; Length 2196;
 Best Local Similarity 98.2%; Pred. No. 0;
 Matches 1660; Conservative 0; Mismatches 1; Indels 30; Gaps 1;

QY 47 CAGGCGACATCTTCAGAAAAAGCATCCCGAGAGAGAGACGAATGTTAAACATCTT 106
 Db 490 CAGGCGACATCTTCAGAAAAAGCATCCCGAGAGAGAGACGAATGTTAAACATCTT 549
 QY 107 AGGTC-----AGCTTAAGCTCTCGAAATTGCT 136
 Db 550 AGGTCAGCTTACAGAGAGAAACCGACACTGTGAGCTTACGCTCTCGAAATTGCT 609
 QY 137 TCTTCAGTGAACCCCGAGAGAGACTGATCACTTCAGTTCTTAAACAAATGCGCCAG 196
 Db 610 TCTTCAGTGAACCCCGAGAGAGACTGATCACTTCAGTTCTTAAACAAATGCGCCAG 669
 QY 197 GTTTGGTACGTTGCGCCGAGTACCATGACATTTCTCAGAGAGAGTGGGCTGTCTGA 256
 Db 670 GTTTGGTACGTTGCGCCGAGTACCATGACATTTCTCAGAGAGAGTGGGCTGTCTGA 729
 QY 257 ACTGTCTCAGAGAGAGCTGTACTGTGGACGTGATGCTGAGAACTACAGTAACTTGTCT 316
 Db 720 ACTGTCTCAGAGAGAGCTGTACTGTGGACGTGATGCTGAGAACTACAGTAACTTGTCT 789
 QY 317 CACTGGAATTTGAGAGTACGATATGAAATTAAGATTTTACTTACAGAAAAAAACATTCATG 376
 Db 790 CACTGGAATTTGAGAGTACGATATGAAATTAAGATTTTACTTACAGAAAAAAACATTCATG 849
 QY 377 AAATTAAGGCTTCCAAAAGAAATTCAGATAGAGAAATTAATCCCTTGGCCGTAACTGGA 436
 Db 850 AAATTAAGGCTTCCAAAAGAAATTCAGATAGAGAAATTAATCCCTTGGCCGTAACTGGA 909
 QY 437 TATGTGAAGTACCGTTGAAGAGACACACGCTCCAGAGAGAGTATGTCAATCAAGTGA 496
 Db 910 TATGTGAAGTACCGTTGAAGAGACACACGCTCCAGAGAGAGTATGTCAATCAAGTGA 969
 QY 497 TCATCAATTAATGTCAAAAAGGCTGTACTAGAGAAAGGACCCCTCTGAACACATCAGA 556
 Db 970 TCATCAATTAATGTCAAAAAGGCTGTACTAGAGAAAGGACCCCTCTGAACACATCAGA 1029
 QY 557 GACATCATAGAGAGAAATCTTTGATGTAGAGACTGTGGAGAGCCCTTATGTCGTGCT 616
 Db 1030 GACATCATAGAGAGAAATCTTTGATGTAGAGACTGTGGAGAGCCCTTATGTCGTGCT 1089
 QY 617 ATCAACTTAATGTCAACATAGAAATCCATCTGTGTGAAGAAACCTTATGAATGAAGAT 676
 Db 1090 ATCAACTTAATGTCAACATAGAAATCCATCTGTGTGAAGAAACCTTATGAATGAAGAT 1149
 QY 677 GTAAGAGAGGCTTCGTTGGGCGAATCAGCTTACTCAATCAAAAAATTCATCTGAGG 736
 Db 1150 GTAAGAGAGGCTTCGTTGGGCGAATCAGCTTACTCAATCAAAAAATTCATCTGAGG 1209
 QY 737 AGAAGCCCTTACGAATGTAAAGACTGTGGAGAGGCTTTTGCATGGGGCTCAAGCTCTGTTA 796
 Db 1210 AGAAGCCCTTACGAATGTAAAGACTGTGGAGAGGCTTTTGCATGGGGCTCAAGCTCTGTTA 1269
 QY 797 TTCATAGAGAGATTCATCTGTGTAAGAAACCTTATGAATGTAAAGACTGTGAGAGGCT 856
 Db 1270 TTCATAGAGAGATTCATCTGTGTAAGAAACCTTATGAATGTAAAGACTGTGAGAGGCT 1329
 QY 857 TTTGGGCGTGTGATGAGCTCACTGACACAGAGATTCACACTGGGAGAGAAAGCTACG 916
 Db 1330 TTTGGGCGTGTGATGAGCTCACTGACACAGAGATTCACACTGGGAGAGAAAGCTACG 1389
 QY 917 AATGCAAGAGCTGTGGAGAGACTTTAGCCGTGTGTATAACTTATCAGACACAGAGAA 976
 Db 1390 AATGCAAGAGCTGTGGAGAGACTTTAGCCGTGTGTATAACTTATCAGACACAGAGAA 1449
 QY 977 TTTCAATGTGGGAGAGAGCTTACAGATGTAAAGACTGTGGAGAGGCTTTTATTTGTGTT 1036
 Db 1450 TTTCAATGTGGGAGAGAGCTTACAGATGTAAAGACTGTGGAGAGGCTTTTATTTGTGTT 1509
 QY 1037 CAAGCCTATTACAGATTAAGAAATTCACACAGGTGAGAAACCTTATGAATGTCAAGAT 1096

ORIGIN

Query Match 80.4%; Score 1619.4; DB 8; Length 2196;
 Best Local Similarity 98.2%; Pred. No. 0;
 Matches 1660; Conservative 0; Mismatches 1; Indels 30; Gaps 1;

47 CAGGCGAGCATCTTCAGAAAAAGCATCCCGAGGAGGAGAGAGCATGGTTAAACATCTT 106
 490 CAGGCGAGCATCTTCAGAAAAAGCATCCCGAGGAGGAGAGAGCATGGTTAAACATCTT 549
 107 AGGTC-----AGCTTAGCCTCTCGAATTTGTCT 136
 550 AGGTGAGCTTTACAGTAGAGAAACGACACTGTGAGCTCTAGCCTCTCGAATTTGTCT 609
 137 TCTTCAATGAGAAACCCCGAGAAAGACTGATCACTTCTTCAATGAGAAACCCCGAG 196
 610 TCTTCAATGAGAAACCCCGAGAAAGACTGATCACTTCTTCAATGAGAAACCCCGAG 669
 197 GTTGTGAGCTTTCGCGGAGCTAGGCAATGACCTTTCTCAGGAGGAGTGGGCTGTCTGA 256
 670 GTTGTGAGCTTTCGCGGAGCTAGGCAATGACCTTTCTCAGGAGGAGTGGGCTGTCTGA 729
 257 ACTCTGCTCAGAGGACCTGTACTGAGGAGCTGATGCTGAGAGAACTACAGTAACTTGTCT 316
 730 ACTCTGCTCAGAGGACCTGTACTGAGGAGCTGATGCTGAGAGAACTACAGTAACTTGTCT 789
 317 CACTGGAATTTGGAGTACAGATATGAAATTAAGATTTCCTACAGAAAAAAACATTATG 376
 790 CACTGGAATTTGGAGTACAGATATGAAATTAAGATTTCCTACAGAAAAAAACATTATG 849
 377 AAATTAAGGCTCTCCAAAAGAAATTCAGATAGAGAAAGTAAATCCCTGGCCCTTACTGGA 436
 850 AAATTAAGGCTCTCCAAAAGAAATTCAGATAGAGAAAGTAAATCCCTGGCCCTTACTGGA 909
 437 TATGTAAGGTAAGCTTTGAAAGACCAAGCGCTCCAGAGGAGTATGTCAATCAGATGA 496
 910 TATGTAAGGTAAGCTTTGAAAGACCAAGCGCTCCAGAGGAGTATGTCAATCAGATGA 969
 497 TCATCAATTAATGTCAAAGGCTGCTACTAGAGAAAGGCAACCCCTCTAGAACATCAGA 556
 970 TCATCAATTAATGTCAAAGGCTGCTACTAGAGAAAGGCAACCCCTCTAGAACATCAGA 1029
 557 GACATCATAGAGAAATTCCTTTGAATGTAAGACCTGTGGGAGGCTTTACTCGTGTCT 616
 1030 GACATCATAGAGAAATTCCTTTGAATGTAAGACCTGTGGGAGGCTTTACTCGTGTCT 1089
 617 ATCAACTTAGTCAATCAGAAATTCATCTGTGAGAAACCTTATGATGTAAGAAAT 676
 1090 ATCAACTTAGTCAATCAGAAATTCATCTGTGAGAAACCTTATGATGTAAGAAAT 1149
 677 GTAAGAGGCTTTCGTTGGGGCAATCAGCTTACTCAACATCAAAAAATTCATCTGGGG 736
 1150 GTAAGAGGCTTTCGTTGGGGCAATCAGCTTACTCAACATCAAAAAATTCATCTGGGG 1209
 737 AGAAGCCCTCAGAAATGTAAGACCTGTGGGAGGCTTTTGAAGGCGCTCAACCCCTGTA 796
 1210 AGAAGCCCTCAGAAATGTAAGACCTGTGGGAGGCTTTTGAAGGCGCTCAACCCCTGTA 1269
 797 TTTCAATAGAGATTTATATCTGTGAAAAACCTTATGATGTAAGAACTGTGAGAAAGGCT 856
 1270 TTTCAATAGAGATTTATATCTGTGAAAAACCTTATGATGTAAGAACTGTGAGAAAGGCT 1329
 857 TTTGGCGGTGTGATGAGTCACTCAGACACAGAGATTTCACACTGGGAGGAAAGACTAG 916
 1330 TTTGGCGGTGTGATGAGTCACTCAGACACAGAGATTTCACACTGGGAGGAAAGACTAG 1389
 917 AATGCAAGAGCTGTGGGAAAGACTTTAGCCGTGTGATTAATCTTATTCAGCAAGAGAA 976
 1330 AATGCAAGAGCTGTGGGAAAGACTTTAGCCGTGTGATTAATCTTATTCAGCAAGAGAA 1449
 977 TTTCAATAGTGGGAGGAAAGCTTTACAGAGTGAAGAACTGTGGGAGGCTTTATTTGTGTT 1036
 1450 TTTCAATAGTGGGAGGAAAGCTTTACAGAGTGAAGAACTGTGGGAGGCTTTATTTGTGTT 1509

1037 CAAAGCTCATTCAGACATTAAGAAATTCACAGAGTGAAGAAACCCATNGATGTCAAGAT 1096
 1510 CAAAGCTCATTCAGACATTAAGAAATTCACAGAGTGAAGAAACCCATNGATGTCAAGAT 1569
 1097 GTGGGAAGGCTTTACTGAGTCAATTTACTTACTCAGCATCAGAAAGATCCACACCGGTG 1156
 1570 GTGGGAAGGCTTTACTGAGTCAATTTACTTACTCAGCATCAGAAAGATCCACACCGGTG 1629
 1157 AGAAGCTTCAGAAATGTAAGAGTGTGGAAAGGCTTTTGTGGGAGTTCAGACCTCGTTA 1216
 1630 AGAAGCTTCAGAAATGTAAGAGTGTGGAAAGGCTTTTGTGGGAGTTCAGACCTCGTTA 1689
 1217 AGCAGAGAGATTCATCAGGCGGAGAGAGCCGCTCAAGTGCACAGAAATGTGGAGGCT 1276
 1690 AGCAGAGAGATTCATCAGGCGGAGAGAGCCGCTCAAGTGCACAGAAATGTGGAGGCT 1749
 1277 TCAATTTGTGCTATCAGCTCAGCAGAGAGATTCACACAGGCGAAACCCCGTATTA 1336
 1750 TCAATTTGTGCTATCAGCTCAGCAGAGAGATTCACACAGGCGAAACCCCGTATTA 1809
 1337 AATGTAAGAGTGTGGGAAAGGCTTTGATTTATGATGAGTCAAGCTCTGTGAACATGAGAA 1396
 1810 AATGTAAGAGTGTGGGAAAGGCTTTGATTTATGATGAGTCAAGCTCTGTGAACATGAGAA 1869
 1397 TTTCAATCCGGGAGTAAACCTTATGAGTGTACAGAAATGTGGGAAAGCTTTAGTCAAGCC 1456
 1870 TTTCAATCCGGGAGTAAACCTTATGAGTGTACAGAAATGTGGGAAAGCTTTAGTCAAGCC 1929
 1457 ATCAGCTTACCAACATCAGAAAGAGGAGGAGTGGGCGGAAATCCATGAAATGTAAGAGT 1516
 1930 ATCAGCTTACCAACATCAGAAAGAGGAGTGGGCGGAAATCCATGAAATGTAAGAGT 1989
 1517 GCGGGAAGCATGTGTAACCACTTAACCATCTCCGAGAAATCAGAGATTCACACACAGTT 1576
 1990 GCGGGAAGCATGTGTAACCACTTAACCATCTCCGAGAAATCAGAGATTCACACACAGTT 2049
 1577 GAAAGGCTTTTGAACGAGTGAAGCCGCTGTATATATGAGTTTGGCTTTTCCACAGTTGT 1636
 2050 GAAAGGCTTTTGAACGAGTGAAGCCGCTGTATATATGAGTTTGGCTTTTCCACAGTTGT 2109
 1637 TACCTGAGTCAACCTGAGTTCAAAAAATATTAATGAAAAATTCAGAAATTAAGAAATTT 1696
 2110 TACCTGAGTCAACCTGAGTTCAAAAAATATTAATGAAAAATTCAGAAATTAAGAAATTT 2169
 1697 TAAAGTCTCAA 1707
 2170 TAAAGTCTCAA 2180

RESULT 14
 AC146884/C
 LOCUS AC146884
 DEFINITION Callithrix jacchus clone CH259-252N7, complete sequence.
 ACCESSION AC146884
 VERSION AC146884.3 GI:68268082
 KEYWORDS HTG;
 SOURCE Callithrix jacchus (white-tufted-ear marmoset)
 ORGANISM Callithrix jacchus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Platyrrhini;
 Callitrichidae; Callitrich.
 1 (bases 1 to 21932)
 Antconellia,A., Ayele,K., Bass,D., Benjamin,B., Bera,J.,
 Blakeley,R.W., Bouffard,G.G., Brinkley,C., Brooks,S., Chu,G.,
 Coleman,H., Engle,J., Fukukenko,T., Gestole,M., Greene,A., Guan,X.,
 Gupta,J., Haghighi,P., Han,J., Hansen,N., Ho,S.-L., Hu,P.,
 Hunter,G., Hurle,B., Idol,J.R., Kwong,P., Laric,P., Larson,S.,
 Lee-Lin,S.-Q., Legaspi,R., Madden,M., Maduro,Q.L., Maduro,V.B.,
 Mullikin,J.C., Oestreicher,J.S., Park,M., Portnoy,M.E., Praead,A.,
 Puri,O., Reddix-Dugue,N., Sante,A., Schandler,K., Schueler,M.G.,
 Stoen,C., Stantropop,S., Stephen,E., Tave,A., Thomas,J.W.,

TITLE Thomas,P.J., Tsipouri,V., Ung,L., Vogt,J.L., Wecherby,K.D.,
 Wither,T.R., Young,A. and Green,E.D.
 JOURNAL NISC Comparative Sequencing Initiative
 REFERENCE Unpublished
 AUTHORS 2 (bases 1 to 219392)
 JOURNAL Direct Submission
 TITLE Submitted (16-OCT-2003) NIH Intramural Sequencing Center, 8717
 Groveont Circle, Galtersburg, MD 20877, USA
 REFERENCE 3 (bases 1 to 219392)
 AUTHORS Green,E.D.
 JOURNAL Direct Submission
 TITLE Submitted (13-FEB-2004) NIH Intramural Sequencing Center, 8717
 Groveont Circle, Galtersburg, MD 20877, USA
 REFERENCE 4 (bases 1 to 219392)
 AUTHORS Green,E.D.
 JOURNAL Direct Submission
 TITLE Submitted (28-JUN-2005) NIH Intramural Sequencing Center, 5625
 Fishers Lane, Rockville, MD 20852, USA
 REFERENCE On Jun 28, 2005 this sequence version replaced gi:42544004.
 JOURNAL
 TITLE
 AUTHORS
 COMMENT
 ----- Genome Center
 Center: NIH Intramural Sequencing Center
 Center code: NISC
 Web site: http://www.nisc.nih.gov
 Contact: nisc_zoo@nigrl.nih.gov
 ----- Project Information
 Center project name: fay
 Center clone name: 252N07

This sequence was finished as follows unless otherwise noted:
 all regions were double-stranded, sequenced with an
 alternate chemistry, or covered by high quality data
 (i.e., phred quality >= 30); an attempt was made to resolve
 all sequencing problems, such as compressions and repeats;
 all regions were covered by at least one plasmid subclone
 or more than one M13 subclone; and the assembly was confirmed
 by restriction digest.

CLONE LENGTH: This sequence represents the entire insert of
 this clone unless otherwise noted. If there are overlapping
 clones, the overlaps are noted in the beginning and end of
 the features section.
 Location/Qualifiers
 1. 219392
 /organism="Callithrix jacchus"
 /mol_type="genomic DNA"
 /db_xref="taxon:9483"
 /clone_id="CH259-252N7"
 /clone_1b="CH259"
 /note="BAC resource: http://bacpac.chori.org/"
 122373..123209
 /note="single clone coverage"
 126183
 /note="low quality single stranded/single chemistry
 region"
 126185
 /note="low quality single stranded/single chemistry
 region"
 197458..197739
 /note="unresolved dinucleotide or trinucleotide repeat"
 200614..200678
 /note="single clone coverage"
 216421..216456
 /note="single clone coverage"
 216694..216853
 /note="single clone coverage"
 217281
 /note="low quality single stranded/single chemistry
 region"
 217283..217285
 /note="low quality single stranded/single chemistry
 region"
 217287..217293

/note="low quality single stranded/single chemistry
 region"
 217317..217319
 /note="low quality single stranded/single chemistry
 region"
 217328..217329
 /note="low quality single stranded/single chemistry
 region"
 217368..217369
 /note="low quality single stranded/single chemistry
 region"
 217392..217393
 /note="low quality single stranded/single chemistry
 region"
 217446
 /note="low quality single stranded/single chemistry
 region"
 217483..217484
 /note="low quality single stranded/single chemistry
 region"
 217495..217497
 /note="low quality single stranded/single chemistry
 region"

ORIGIN

Query Match 72.9%; Score 1469.6; DB 8; Length 219392;
 Best Local Similarity 88.2%; Pred. No. 0;
 Matches 1674; Conservative 0; Mismatches 199; Indels 25; Gaps 6;

114 CTTGAGGAGAGTGGGAGTCTGTCTGAACTCTGCTCAAGAGGAGCTGTACTGGAGCTGATG 171
 156952 CTTGAGGAGAGTGGGAGTCTGTCTGAACTCTGCTCAAGAGGAGCTGTACTGGAGCTGATG 156893
 172 TTCAAGTTCTAAACAATGAGCCAGAGGTTTGTAAGCTTGCAGAGTACCTAGACTTT 231
 156892 TTCAATTCATAACAATGAGCCAGAGGTTTGTAAGCTTGCAGAGTACCTAGACTTT 156833
 232 TCTCAGAGAGAGTGGGAGTCTGTCTGAACTCTGCTCAAGAGGAGCTGTACTGGAGCTGATG 291
 156832 TCTCAGAGAGAGTGGGAGTCTGTCTGAACTCTGCTCAAGAGGAGCTGTACTGGAGCTGATG 156773
 292 CTGAGAGAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 351
 156772 CTGAGAGAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 156713
 352 TTACCTACAGAAAAAACAATCATGAAATTAAGGCTCCAAAGAGAAATTCAGATGAGAGA 411
 156712 TTACCTACAGAAAAAACAATCATGAAATTAAGGCTCCAAAGAGAAATTCAGATGAGAGA 156653
 412 AGTAAATCCCTTGGCCGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 471
 156652 AGTAAATCCCTTGGCCGATGATGATGATGATGATGATGATGATGATGATGATGATG 156593
 472 AAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 531
 156592 AAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 156533
 532 GGCAGCCCTCTAGAACATCATGAGACATCATTAAGAGAAATTCCTTGAATGTAAGAGC 591
 156532 GGCAGCCCTCTAGAACATCATGAGACATCATTAAGAGAAATTCCTTGAATGTAAGAGC 156473
 592 TGTGGAGAGGCTTTAGTGTGCTGATGATGATGATGATGATGATGATGATGATGATGATG 651
 156472 TGTGGAGAGGCTTTAGTGTGCTGATGATGATGATGATGATGATGATGATGATGATGATG 156413
 652 GAGAAACCTTAGAATGTAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAG 711
 156412 GAGAAACCTTAGAATGTAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAG 156353
 712 CAACATCAAAAAATTCATAGTGGGAGAGCCCTTACGAATGTAAGAGCTGTGGAGAGGCT 771
 156352 CAACATCAAAAAATTCATAGTGGGAGAGCCCTTACGAATGTAAGAGCTGTGGAGAGGCT 156293

OY	772	TTTTGATGGGGCTCAAGCCCTGGTATTCTAATAAGAAATTCTATCTGTGAAAAACCTAT	831
Db	156292	TTTCCATGGGGCTCAAGCCCTGGTATTCTAATAAGAAATTCTATCTGGGGAATCTTAC	156233
OY	832	GAATGTAAAGACTGTGGAAGAGCCCTTCCGCGGTGTAAGAGCTCACTCGACACGAGA	891
Db	156232	GAATGCAAAGACTGTGGAAGAGCCCTTCCGCGGTGCGCAAGATTCCGCGACGACCA----	156177
OY	892	TTCCACACTGGGGAAGAACTACGAATGCAAAGACTGTGGGAAGACTTTTACCGGTGTG	951
Db	156176	-----TGTGCAGACTCAAGGTCAAAAGAC--TGGAAGACTTTTATGCC--GGG	156132
OY	952	TATAAATTAATTCAGACAAAGAAATTCTATGTGGGGAAGAGCCCTTACAGAGTAAAGAC	1011
Db	156131	TATAACATTAATTCAGACAAAGAAATTCAAGTGGGAAGAGCCCTTACAAATGTAAAGAC	156072
OY	1012	TGTGGGAAGGCTTTTATTGTGGTTACGCTCATTCAGACTAAAGAAATTCACACAGGT	1071
Db	156071	TGTGGGAAGGCTTTTATTGTGTGCTTAAGCTTCAATTTAGATTAAGAAATTCACACAGGT	156012
OY	1072	GAGAAACCCATGAATGTCAAGAAATGTGGGAAGCCCTTTACTGAGTCAATTACTTACT	1131
Db	156011	GAGAAACCCATGAATGTCAAGAAATGTGGGAAGCCCTTCACTCAATTAATGACTTAACT	155922
OY	1132	CAGCATCAGAAGATCAACCGGTGAAGAGCTTCAAGATGTAAAGATGTGGGAAGGCC	1191
Db	155951	CGGCATCAGAAGATCAATCATCAAGTGAAGAGCCCAAGAGTAAAGAGTGCAGAAAGGCC	155892
OY	1192	TTTGGCTGGGGTTTCGAGCCTCGTTAGCAAGAGGATCATACGGGCGAAGAGCCGTAC	1251
Db	155891	TTTGGCTGGGGTTTCAGTCTTGTGAAGCAAGAGGATCATCAATGGGCGAAGAGCCGTAC	155832
OY	1252	AAGTGCACGAATGTGGGAAGGCTTCAATTTGGCTATATCACTCACTCAGCAGAGAGA	1311
Db	155831	AAATGCACGAATGTGGGAAGGCTTCAACTGTGGCTATATCTTAATCTCAACAGAGAGA	155772
OY	1312	ATCCAACAGCGGAAACCCCGTATAAATGTAAAGAGTGTGGGAAGGCTTCAATTAAGAA	1371
Db	155771	ATCCAACAGAGGAAACCCGTGTCAAAATGTAAAGAGTGTGGGAAGGCTTCAATTAAGGC	155712
OY	1372	TGGAACCTCGTGAACATATAGAGAAATTCATACCGGGGTGAACCTTATGGGTATACGAA	1431
Db	155711	TTGACCTCGGTGAACACAGAGAAATTCATACCAAGGTGAAGGCCCTACGGGTATACAGAA	155652
OY	1432	TGTGGGAAGAGCTTTAGTCAAGGCGCATCAGCTTACCAACATCAGAAAAACGACAGTGGG	1491
Db	155651	TGTGAAGAGAGCTTTTAGTCAATGCCATCAGCTTACCAAGCACAGAAAAACACAGTGGG	155592
OY	1492	GCGAAATCTCAGATGTAAAGAGTGTGGGAAGGAGTGAACCACTTAACCATCTCCGA	1551
Db	155591	CGAAATCTCTAATAATGTAAAGAAATGTGGGAAGGCGTGTCAACAGCTTAACCTCTCCGA	155322
OY	1552	GAACTTCAGAGATCCACAACAGTTGAAGACCTTTTGAACGAGATAGCCGCTCGTATTC	1611
Db	155531	GAACTTCAGAGATCCACAACAGTTGATAGCCCTTTTGAACCTCAGCCGCTCGTATATC	155472
OY	1612	TATGGTTTGGCTTTTCCACAGTTTGTAACTGACGTCAGACTCAGCTGCTCAAAAAATTAAT	1671
Db	155471	CGTGAATTCGCTCTCCACATTTTGTTAACCTGACGTCAACTGCACTTCAAAAAATTAAT	155412
OY	1672	GGAATAATTCAGAAAAATTAAGAAATTTTAAGTCTCAAAATGTGTGACCCCTTGTGATAGCTG	1731
Db	155411	GGAAATTCAGAAATA---AATTGTAAGTTCAAAATGTGTGACCCCTTGTGATAGCGGG	155355
OY	1732	ATGAATCTCTCGCTGTCCGCTCCAGCCGGCCGGGGAATGTAGTCAATCCCTTGTGTCAAG	1791
Db	155354	ATGAATCTCTTGTGTCTCTGTCGCAATCAGCCAGGAATGTGACTTGCACTTGTGTCAAG	155295
OY	1792	CACATTCACGCTGTATATCCACACCCACCCCTGCTAGTGAATTAAGACGCTTGGTGATC	1851
Db	155294	CACATTCACGCTGTATATG--CTGCAATCTCTGTATGTCATTAATAGCTGCTTGGTGATC	155236
OY	1852	AGATCAACTATCCAGCATCAGAGTGCCTGTGCGCAAGTATGCTCTCACTTGTGCTTAACAG	1911

Db	155235	AGATGACCTGTACGGCATTACAGTGCCTTGTCCTAAGTAATCTTTACTTCACTTAAACG	155176
Qy	1912	TGGCCCCAGAGAGCAGAAGTAGTGAATGCTGTGATTCGGATATGCCAAAGAGAACACA	1971
Db	155175	TGGCCCCAAAGAGCAAGAGTAGTGAATGCTGTGATTCGGATATTCGGATATTCGGAAAGAGAACACA	155116
Qy	1972	AAGTCTTCTTTTAAATGAAAAAGGTGAAGTTCTCAA	2009
Db	155115	AAGTCTTCTTTTAAATGAAAAAGGTGAAGTTCTCAA	155078
RESULT 15			
LOCUS	CQ718116	1420 bp	DNA linear PAT 03-FEB-2004
DEFINITION	Sequence 4050 from Patent WO02068579.		
ACCESSION	CQ718116		
VERSION	CQ718116.1 GI:42278973		
KEYWORDS			
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;		
TITLE	Hominidae; Homo.		
JOURNAL	Venter, C.J., Adams, M.C., Li, P.W. and Myers, E.W. Kites, such as nucleic acid arrays, comprising a majority of humanecons or transcripts, for detecting expression and other uses thereof Patent: WO 02068579-A 4050 06-SEP-2002; PE Corporation (NY) (US); Location/Qualifiers 1..1420 /organism="Homo sapiens" /mol_type="unassigned DNA" /db_xref="taxon:9606"		
FEATURES			
source			
ORIGIN			
Query Match	69.9%; Score 1408.8; DB 6; Length 1420;		
Best Local Similarity	99.9%; Pred. No. 0;		
Matches 1410; Conservative	0; Mismatches 2; Indels 0; Gaps 0;		
Qy	114	CTTAGCCCTCGGAATTTGTTCTTTCAGTGGAAACCCCGAAGACTGATCACTTCTT	173
Db	1	CTTAGCCCTCGGAATTTGTTCTTTCAGTGGAAACCCCGAAGACTGATCACTTCTT	60
Qy	174	CAGTTCTAAAACAATGCGCCAGGGTTTGGTACGTTCCGCCAGTACCATGACCTTTTC	223
Db	61	CAGTTCTAAAACAATGCGCCAGGGTTTGGTACGTTCCGCCAGTACCATGACCTTTTC	120
Qy	234	TCAGAGAGAGTGGGCTGTCTGAACTCTGCTCAGAGGGACTGTACTGGGACGTGATGCT	293
Db	121	TCAGAGAGAGTGGGCTGTCTGAACTCTGCTCAGAGGGACTGTACTGGGACGTGATGCT	180
Qy	294	GGAGAACTACAGTAACTTGGTCTCCTGGATTGGAGTCAAGATATGAAATAAGAGTTT	353
Db	181	GGAGAACTACAGTAACTTGGTCTCCTGGATTGGAGTCAAGATATGAAATAAGAGTTT	240
Qy	354	ACCTTACAGAAAAAACAATTCATGAAATAAGGCTTCCAAAGAAGATTGATGAAAGA	413
Db	241	ACCTTACAGAAAAAACAATTCATGAAATAAGGCTTCCAAAGAAGATTGATGAAAGA	300
Qy	414	TAAATCCCTTGGCCGTTAACTGGAATATGGAAGGTACGTTGAAAGACACAGCGCTCCAG	473
Db	301	TAAATCCCTTGGCCGTTAACTGGAATATGGAAGGTACGTTGAAAGACACAGCGCTCCAG	360
Qy	474	AGGAGGTATGTCAATCAGATGATCATCAATTTATGTCAAAAGCGCTGCTACTAGAGAAG	533
Db	361	AGGAGGTATGTCAATCAGATGATCATCAATTTATGTCAAAAGCGCTGCTACTAGAGAAG	420
Qy	534	CACCCTCTTGAACAACATCAGAGCATCATTAAGAGAAATTCCTTTGAATGTGAAGACTG	593
Db	421	CACCCTCTTGAACAACATCAGAGCATCATTAAGAGAAATTCCTTTGAATGTGAAGACTG	480

```
OY 594 TGGGAAGGCTTTAGTCGTGCTATCACTTAGTCAATCAGAAAAATCCATATCTGTGA 653
| | | | |
DB 481 TGGGAAGGCTTTAGTCGTGCTATCACTTAGTCAATCAGAAAAATCCATATCTGTGA 540
| | | | |
OY 654 GAAACCTTATGAATGTAAGAAATGTAAGAAAGCTTCCGTTGGGGCAATCAGCTTACTCA 713
| | | | |
DB 541 GAAACCTTATGAATGTAAGAAATGTAAGAAAGCTTCCGTTGGGGCAATCAGCTTACTCA 600
| | | | |
OY 714 ACATCAAAAAATTTCACTCTGGGGAAGGCCCTACGAATGTAAGAACTGTGGGAAGCTTT 773
| | | | |
DB 601 ACATCAAAAAATTTCACTCTGGGGAAGGCCCTACGAATGTAAGAACTGTGGGAAGCTTT 660
| | | | |
OY 774 TCGATGGGGCTCAAGGCTCGTTATTATTAAGAGATTCTACTGTGAAAAAACCCTATGA 833
| | | | |
DB 661 TCGATGGGGCTCAAGGCTCGTTATTCTAAGAGATTCTACTGTGAAAAAACCCTATGA 720
| | | | |
OY 834 ATGTAAAGA CTGTGAAAAGGCTTTCCGCGTGTGATGAGCTCACTCAGACCAAGATTT 893
| | | | |
DB 721 ATGTAAAGA CTGTGAAAAGGCTTTCCGCGTGTGATGAGCTCACTCAGACCAAGATTT 780
| | | | |
OY 894 CCAACTGGGGGAAGAAAGACTAAGAAATGCAAGAACTGTGGGAAGAACCTTTAGCCGTGTGA 953
| | | | |
DB 781 CCAACTGGGGGAAGAAAGACTAAGAAATGCAAGAACTGTGGGAAGAACCTTTAGCCGTGTGA 840
| | | | |
OY 954 TAACTTATTCAGCAAGAGAAATTCATAGTGGGGAAGACCTTACGAGTGTAAAGACTG 1013
| | | | |
DB 841 TAACTTATTCAGCAAGAGAAATTCATAGTGGGGAAGACCTTACGAGTGTAAAGACTG 900
| | | | |
OY 1014 TGGGAAGGCTTTTATTGTTGTTCAAGCCTCATTCAGCATMAAAGAAATTCACACAGTGA 1073
| | | | |
DB 901 TGGGAAGGCTTTTATTGTTGTTCAAGCCTCATTCAGCATMAAAGAAATTCACACAGTGA 960
| | | | |
OY 1074 GAAACCTTATGAATGTAAGAAATGTAAGAAAGGCTTTACTCGAGTCAATTAACCTTACTCA 1133
| | | | |
DB 961 GAAACCTTATGAATGTAAGAAATGTAAGAAAGGCTTTACTCGAGTCAATTAACCTTACTCA 1020
| | | | |
OY 1134 GCATCAGAGATTCACACCGGTGAGAAGCTTCAGAAATGTAAGAGTGTGGGAAGGCTTT 1193
| | | | |
DB 1021 GCATCAGAGATTCACACCGGTGAGAAGCTTCAGAAATGTAAGAGTGTGGGAAGGCTTT 1080
| | | | |
OY 1194 TGGCTGGGGTTCGAGCTCGTTAAGCAGAGAGATACATACGGGCGAAGGCCGTACAA 1253
| | | | |
DB 1081 TGGCTGGGGTTCGAGCTCGTTAAGCAGAGAGATACATACGGGCGAAGGCCGTACAA 1140
| | | | |
OY 1254 GTGCACAGAAATGTGGGAAGGCTTCAATTGTGCTATCACTCACTCAGACGAGAGAAAT 1313
| | | | |
DB 1141 GTGCACAGAAATGTGGGAAGGCTTCAATTGTGCTATCACTCACTCAGACGAGAGAAAT 1200
| | | | |
OY 1314 CCAACAGAGGCAAAACCCGTAATAATGTAAGAGTGTGGGAAGGCTTTCAATTTATGATC 1373
| | | | |
DB 1201 CCAACAGAGGCAAAACCCGTAATAATGTAAGAGTGTGGGAAGGCTTTCAATTTATGATC 1260
| | | | |
OY 1374 GAGCCTGTGAAACATGAGAGAAATTCATACCGGGGTGAAAACCTATGGGTGTACAGAAATG 1433
| | | | |
DB 1261 GAGCCTGTGAAACATGAGAGAAATTCATACCGGGGTGAAAACCTATGGGTGTACAGAAATG 1320
| | | | |
OY 1434 TGGGAAGAGCTTATGTCACGGCCATCATGACTTACACAAATCAGAAAAACGACAGTGGGGC 1493
| | | | |
DB 1321 TGGGAAGAGCTTATGTCACGGCCATCATGACTTACACAAATCAGAAAAACGACAGTGGGGC 1380
| | | | |
OY 1494 GAAATCTACGAATGTAAAGAGTGGGGAAGG 1525
| | | | |
DB 1381 GAAATCTACGAATGTAAAGAGTGGGGAAG 1412
| | | | |
```

Search completed: February 6, 2006, 19:16:50
Job time : 10240 secs